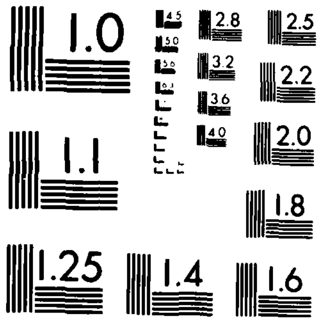
AD-A152 229 THE DESIGN OF AN EXPERIMENT TO EXAMINE REPAIR PROCESS ERRORS OF MILITARY VEHICLE MECHANICS(U) FLORIDA INST OF TECH MELBOURNE D T CLEMENTS 07 DEC 84 1/2

UNCLASSIFIED F/G 5/9 NL


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

# REPORT DOCUMENTATION PAGE

| | |
|---|---|
| 1a REPORT SECURITY CLASSIFICATION Unclassified | 1b. RESTRICTIVE MARKINGS |
| 2a SECURITY CLASSIFICATION AUTHORITY N/A | 3. DISTRIBUTION/AVAILABILITY OF REPORT |
| 2b DECLASSIFICATION/DOWNGRADING SCHEDULE N/A | |
| 4 PERFORMING ORGANIZATION REPORT NUMBER(S) | 5 MONITORING ORGANIZATION REPORT NUMBER(S) US Army Concepts Analysis Agency Bethesda, Md. |
| 6a NAME OF PERFORMING ORGANIZATION Florida Inst. of Tech. Melbourne, Fl. — 6b OFFICE SYMBOL (If applicable) | 7a. NAME OF MONITORING ORGANIZATION US Army Concepts Analysis Agency Bethesda, Md. |
| 6c ADDRESS (City, State, and ZIP Code) Florida Institute of Technology 150 West University Blvd. Melbourne Fl. 32901 | 7b. ADDRESS (City, State, and ZIP Code) |
| 8a NAME OF FUNDING/SPONSORING ORGANIZATION — 8b. OFFICE SYMBOL (If applicable) | 9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER |
| 8c ADDRESS (City, State, and ZIP Code) | 10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. / PROJECT NO / TASK NO. / WORK UNIT ACCESSION NO. |

DTIC ELECTE APR 5 1985 S A

# AD-A152 229

11 TITLE (Include Security Classification)
The Design of an Experiment to Examine Repair Process Errors of Military Vehicle Mechanics (Unclassified)

12 PERSONAL AUTHOR(S)
Maj. Denis Thomas Clements

| 13a TYPE OF REPORT | 13b. TIME COVERED | 14. DATE OF REPORT (Year, Month, Day) | 15. PAGE COUNT |
|---|---|---|---|
| Research Thesis | FROM ______ TO ______ | 1984 December 7 | 117 |

16 SUPPLEMENTARY NOTATION This thesis was developed upon a research proposal provided to the author by the US Army Concepts Analysis Agency, a Field Operating Agency of the Director of the Army Staff.

| 17 COSATI CODES | | | 18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) |
|---|---|---|---|
| FIELD | GROUP | SUB-GROUP | Vehicle Mechanics, Diagnosis of Faults, Vehicle Repair, Mechanic's Errors, Statistical Analysis, Canonical Correlation, Human Factors, Ops Research |

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

This thesis develops a method and concept for analysis of errors made by US Army vehicle mechanics. A process model is developed to describe the heirarchy of actions accomplished by soldier mechanics to complete a diagnosis and repair effort on a disabled vehicle. From this process model an error classification scheme is developed. An error list is derived from the error classification scheme and used in combination with a list of factors that contribute to soldier mechanic's errors to determine shortcomings in the US Army system that selects, trains, employs, and provisions soldier mechanics. An experiment is developed which allows non-intervening observers to collect information regarding the incidence of error types with their associated contributing factors. This information is used in a statistical analysis. The analytical method used is canonical correlation. Canonical correlation analysis produces a rank ordering and

| | |
|---|---|
| 20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS | 21. ABSTRACT SECURITY CLASSIFICATION Unclassified |
| 22a NAME OF RESPONSIBLE INDIVIDUAL Maj. Denis Thomas Clements | 22b TELEPHONE (Include Area Code) 314-843-9758 — 22c OFFICE SYMBOL |

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

DTIC FILE COPY

85 03 15 143

relative scaling of the factors that contribute to soldier mechanic's errors. This analytical result may then be used by top-level US Army decision-makers when deciding the allocation of research and development funds to reduce the frequency of errors made by soldier mechanics, thus improving the overall effectiveness of the US Army maintenance effort.

THE DESIGN OF AN EXPERIMENT TO EXAMINE REPAIR PROCESS

ERRORS OF MILITARY VEHICLE MECHANICS

---

A Thesis

Presented to

the Faculty of the Graduate School

Florida Institute of Technology

---

In Partial Fulfillment

of the Requirements for the Degree of

Master of Science



---

by

Denis Thomas Clements

December 1984

# THE DESIGN OF AN EXPERIMENT TO EXAMINE REPAIR PROCESS ERRORS OF MILITARY VEHICLE MECHANICS


BY

DENIS THOMAS CLEMENTS


A Thesis submitted in partial fulfillment

of the requirements for the degree of

Master of Science


FLORIDA INSTITUTE OF TECHNOLOGY


1984

Approved by ____________ Date 12/4/84


Department of Mathematical Sciences

and Computer Science

## ABSTRACT

This thesis develops a method and concept for analysis of errors made by US Army vehicle mechanics. A process model is developed to describe the heirarchy of actions accomplished by soldier mechanics to complete a diagnosis and repair effort on a disabled vehicle. From this process model an error classification scheme is developed. An error list is derived from the error classification scheme and used in combination with a list of factors that contribute to soldier mechanic's errors to determine shortcomings in the US Army system that selects, trains, employs, and provisions soldier mechanics. An experiment is developed which allows non-intervening observers to collect information regarding the incidence of error types with their associated contributing factors. This information is used in a statistical analysis. The analytical method used is canonical correlation. Canonical correlation analysis produces a rank ordering and relative scaling of the factors that contribute to soldier mechanic's errors. This analytical result may then be used by top-level US Army decision-makers when deciding the allocation of research and development funds to reduce the frequency of errors made by soldier mechanics, thus

improving the overall effectiveness of the US Army maintenance effort. This thesis was developed upon a research proposal provided to the author by the US Army Concepts Analysis Agency, a Field Operating Agency of the Director of the Army Staff.

ACKNOWLEDGEMENTS

The following acknowledgements are made for two reasons. First to extend my appreciation to these individuals who took the time to discuss various issues within the field of mechanical maintenance. And secondly, to provide a record of the organizations that provided the most insight in the formulation of this thesis. The second point must be qualified by stating that these individuals should in no way be associated with any shortcomings that may exist in this thesis, those shortcomings are strictly the resposibility of the author.

Dr. Richard Kern, Dr. Melissa Berkowitz, and Dr. Robert Wisher of the US Army Institute of the Behavioral and Social Sciences provided assistance regarding many aspects of the soldier mechanic. These included use of printed manuals, management of maintenance operations and maintenance personnel, the development of sophisticated electronic aids for mechanics, and most importantly, the problems involved in conducting experiments in US Army motor pools. Dr. Kern provided considerable insight into both the thesis concept and the idea of using non-intervening observers.

Mr. Cloyd Senn, Mr Harold Oliver, Major David Jones, and Major John

Williams all of the US Army Ordnance Center and School, provided information regarding the training of soldier mechanics. They also provided direct comment on the error types and contributing factors discussed in this thesis.

Several other individuals provided comments regarding current efforts in areas directly related to this thesis. Mr. Tod Stephenson of the US Army Test and Evaluation Command discussed vehicle repairability, Major Tom McGrann of the US Army Logistics Center discussed maintenance manpower requirements. Mr. Ron Beehler, also of the Logistics Center, discussed the Standard Army Maintenance System. Mr. Elmer Williams, a regional maintenance director with United Parcel Service's Atlanta, Ga. office discussed all aspects of how a major commercial company addresses the concern for improvements in efficiency and productivity in vehicle maintenance.

A final thank you is reserved for the many professors and other personnel of the Florida Institute of Technology, for their kind assistance, with a special thank you to Dr. Frederick Buoni for his untiring efforts in guiding the development of this thesis, and for the use of his Macintosh personal computer on which this thesis was typed and printed.

TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# Chapter 1

## PROLOGUE

The US Army loses large sums of money due to the errors made by soldier mechanics in their daily activities of diagnosing and repairing malfunctioning vehicles. It spends large sums of money to improve the *selection, training, and employment of these soldier mechanics*. It also spends a great deal on the test, measurement, and diagnostic equipment used by these soldier mechanics, and on the enhanced maintainability of the vehicles to be serviced or repaired by the mechanics.

It would be desirable to optimize the allocation of this very large amount of money within the Army system. This must be done to focus on the weaknesses most highly correlated with the errors made by soldier mechanics.

*It is the task of the author of this thesis to develop a method to* correlate soldier mechanic's errors and the factors which contribute to those errors. The development goes a step further than this. An argument is also developed to link the factors that contribute to soldier mechanic's errors and the elements of the Army system that the contributing factors represent. In this way systemic weaknesses can be associated with mechanic's errors. The analysis not only leads to the identification of systemic weaknesses but also to a method of ordering the weaknesses. Consequently, ordering allows optimization of monetary-resource allocation to most effectively address the conditions that lead to soldier mechanic errors. The concept of examining human errors and drawing statistical inferences from the examination is complex.

Operations Research brings quantitative science to bear on complex problems. Operations Researchers employ a wide variety of mathematical tools to solve these complex problems.

The use of a designed experiment allows the statistical analysis of a problem in an active manner, a manner that very often exceeds the power of statistical analysis of passively collected data. How this ties into the problem at hand, that of determining the factors that contribute to the soldier mechanic's inability to satisfactorily diagnose and repair equipment failures, is not obvious. It has been an accepted fact that apprentice vehicle mechanics often fail to correctly diagnose the cause of vehicle breakdowns. This human failure results in lengthy repair times, incomplete or faulty repairs which lead to subsequent breakdowns, and excessive parts consumption through the random "trial and error" replacement of parts to achieve a repair. When it is said that this is an accepted fact, this is to say that passive data collection efforts have revealed a serious problem. The problem is the large percentage of repair parts that are being discarded as bad, when in fact they are good, coupled with the high fraction of repair time that is a consequence of this erroneous repair effort.

The problem of mechanic's errors has been revealed by the analysis of passively-collected data, and through the effects of errors, such as lost labor hours, wasted dollars, and lost vehicle time. In spite of this, no insight has been gained into the specific causes of the problem.

Currently, monetary resources are allocated to those elements of the Army system that have been shown to influence the mechanics behavior. Very often the argument to fund improvement programs is driven by apparent opportunities for improvement, and not because the element of the Army system addressed has been proven to be a leading cause of mechanic's errors. In this way, money is allocated to virtually all system elements with no certainty that the culprit elements are being allocated an optimal share of the monetary resource.

## PROBLEM BACKGROUND

The effort to solve the problem of poor diagnostic and poor repair-performance has thus far taken the path of examining each element of the institution of the US Army which has the potential for being a contributor to a mechanic's poor performance. This effort has included constant re-appraisal and improvement of the means used to select, train, employ, and provision the soldier mechanic, as well as improvement in the repairability of the vehicles. The problem-solving effort has not provided a means to isolate and evaluate proven weakness. Consequently, it is difficult to prove that success has been acheived. The reason for this situation is no more clear than is the reason for the mechanic's errors. Without means to identify and isolate causes, there are no means to identify successes.

How does one know that a training program has done a better or worse job in influencing the results of a mechanic's work? Using training as an example, the student soldier can be tested during and after his Initial Entry Training to discover the effects of different training approaches and their relative worth. This is done constantly with great success. But how is one to measure this same worth in the workplace when all of the other influences simultaneously impact on the mechanic's success, or lack of success in doing his work?

The problem of trying to isolate the factors that contribute to human error is not unique to vehicle mechanics in the US Army, it is a problem that arises wherever an effort is mounted to reduce the errors made by skilled humans working in a complex environment.

How can an individual's or a group's efforts be evaluated to determine ways to improve productivity? Assuming that there is some level of dissatisfaction with the status quo, and indeed there is in the case at hand, where are the causes of productivity reduction found? The workplace is the location in which to look. Factor's that contribute to mechanic's errors can be spotted by trained observers working alongside the mechanic in his shop or field location.

## REVIEW OF THE LITERATURE

This thesis was developed from a research proposal provided to the author by the US Army Concepts Analysis Agency. The research proposal discussed the possible connection between poor diagnostic and repair skills. and excess supply actions. The research proposal went on to recommend examination of commercial fleet operators, changes to mechanic selection criteria, lengthening of mechanic training, expansion of mechanic training curricula, evaluation of mechanic diagnostic skill, and finally, changes to existing maintenance management policy. All of these subjects have been examined by the author.

The Brown Board's study of US Army Logistics in 1966 revealed in detail that as many as seven out of ten mechanics failed to pass structured tests of their level of ability in the application of diagnostic skill, US Department of the Army [66]. A test conducted at Ft. Carson, Colorado in 1974, showed that thirty-five percent of the generators, voltage regulators, alternators, distributors, and starter motors turned in as unserviceable were actually in working condition, US Department of the Army [74]. Dressel and Sheilds, of the US Army Research Institute, reported more detailed results in 1979. Their study revealed that thirty percent of parts were replaced in error, consuming thirty-two percent of the mechanic's productive time and accounting for thirty-two percent of vehicle downtime, Dressel and Shields [79]. There is, without question, a problem of considerable proportion. This false replacement is a symptom. The symptom points to the mechanic, and this suggests that the system that put the mechanic there, with the skill, or lack thereof, must be examined.

Solution efforts abound, the Improved Technical Documentation and

Training program put the ball in the lap of the hardware proponent and the contractor/supplier, Buchan and Knutson [77]. The cognitive style of the potential military member was examined to determine if new screening methods would help, Federico [78]. Test, Measurement, and Diagnostic Equipment, along with Automatic Test Equipment, Built-In Test Equipment, and Simplified Test Equipment, have all drawn much attention, Hackenbruch [82], Pantiskas [81], Sweeny [81]. The maintenance management system has received much study, as well as a facelift with the advent of the Standard Army Maintenance System, Harris [84]. The personnel system has been revised both vertically and horizontally by the introduction of new military specialties and a new distribution of responsibilities, US Deapartment of the Army [84], Staff [80], Staff [78]. Training continues to be revised in light of all of the above, as well as for incorporating new directions in its own field, US Department of the Army [84]. Much effort is being directed at the problem of mechanic's errors but it appears to persist. What of this human skill of fault diagnosis? The field of Artificial Intelligence has forced reassessment of many intellectual skills, fault diagnosis among them. Stefik explains that diagnosis is the process of fault finding in a system, based on the interpretation of potentially noisy data. A diagnostician must understand the system organization and the relationships and interactions between sub-systems. Stefik goes on to list the difficulties that attend a diagnostic effort, including faults being masked by ambiguous symptoms, faults that are intermittent, and faulty diagnostic equipment, to name a few, Stefik [82].

The errors that humans commit in the fault diagnosis effort have received attention. Rouse cites Norman in saying that errors are of two types, "slips and mistakes". He goes on to say that the mistakes are the errors that can be dealt with, because, once identified, the mistake, due to the incorrect application of knowledge or lack of knowledge, can be fixed, Rouse and Rouse [83]. Janis and Mann say that the causes of many errors go beyond the factors of selection, training, and supervision, to the psychological state of the human, Janis and Mann [77]. Also, there is a limit to the human's ability to handle complex systems, when performing a diagnostic task, Wohl [83].

Much scientific effort is being brought to bear on the human and his abilities to determine what can be expected from a person. Cognitive style has indeed proven to be a factor in successful diagnostic work on mechanical systems, Rouse and Rouse [79], Rouse and Rouse [82]. Fault diagnosis includes human decision-making, and successful decision-making in diagnostic work is affected by problem size, interconnectedness of parts, and environmental stress. These factors, when present, may cause poor decision choices that would not be made otherwise, Rouse [79]. How humans cope with stress, and how information can be presented to aid in diminishing the impact of stress, has received much attention, Sage [81].

These findings tell us to look at the error first, then find the cause. Human-error analysis can lead to refined systems. Overall measures of diagnostic and repair faults, such as poor operational readiness rates or high false replacement rates, only indicate unsatisfactory performance. These results provide no evidence concerning the task performance process or concerning the individual's understanding of that process, Johnson and Rouse [82]. The human must be studied while controlling factors such as faulty diagnostic equipment and incorrect manuals. By consequence, if the diagnosis and repair effort is erroneous, it is the mechanic's error. If these errors can then be categorized to reflect a correlation between the error and the failure in the system that selects, trains, employs, and provisions the mechanic, it may be possible to prevent the error, Rouse and Rouse [83]. Rouse goes on to demonstrate that error detection must go on in the real environment, with the real players present.

## PROBLEM STATEMENT

The author of this thesis develops a concept to determine whether or not a correlation exists between mechanic's errors and factors that contribute to those errors. *The errors will be categorized by the use of a* process model which will also be developed in this thesis. The contributing factors will be categorized by the element of the Army system that is responsible for controlling the particular contributing factor. As an example, a contributing factor such as lack of basic knowledge concerning mechanical function of the hydraulic brake system of a vehicle could be attributed to lack of adequate coverage of this subject in the Initial Entry Training, suggesting that the Army system element of Institutional Training should be examined. If this example holds for many mechanics and Institutional Training is found to be highly correlated with the errors that mechanics make in general, then monetary resources should be allocated to correct this problem as opposed to elements of the Army system that have a lower correlation with mechanic's errors.

The author then discusses an experimental hypothesis and method of statistical analysis that can be used to detect the above described correlation if it exists. And further, the analysis of the experiment result will allow the rank ordering of systemic weaknesses and their respective contribution to the total of mechanic's errors. This scheme can provide a decision maker with the ability to allocate monetary resources optimally to those system weaknesses that are most responsible for mechanic's errors.

Also within the framework of this experiment design and analysis, will be discussed a second level of the experiment to detect the possible aggravation of the mechanic's error rate when he moves from a garrison environment to a field environment.

The final topic addressed is how data to be analyzed may be collected. This discussion includes the development of the sample, estimates of personnel needed to conduct the experiment and the data-set collection effort, and the potential cost of such an experiment.

## THESIS ORGANIZATION

This thesis is organized to address four issues. The first is the development of an error classification scheme from a process model. This scheme follows a method proposed by Rouse. The second is the development of a list of contributing factors. This list is developed by examining the way a systemic weakness exhibits itself in the workplace as a factor contributing to human error. These first two issues are addressed in Chapter Three.

The development of an experimental hypothesis and the statistical methods of testing it constitute the third issue discussed. The companion issue of designing an experimental framework within which to test the hypothesis is the final issue taken up. These issues are discussed in Chapter Four.

Chapter Five outlines a proposal for the conduct of the experiment. Included are the sample selection, personnel requirements, and possible costs. This last chapter examines the candidate mechanic specialties and the specialties and number of people needed to supervise the experiment, collect the data, and analyze the data.

Prior to embarking on this effort, Chapter Two discusses in general terms, Rouse's method of analyzing and classifying human error, Rouse and Rouse [83]. Also discussed is Van Eekhout's work that includes the idea of correlating human error with contributing factors. The first portion of Chapter Two lays the groundwork for Chapter Three. The second portion of Chapter Two discusses some fundamental ideas of the design and analysis of experiments. This discussion provides the basis for Chapter Four.

# Chapter 2

## CLASSIFICATION AND ANALYSIS OF HUMAN ERROR

The classification and analysis of human error is contained in the study of human factors. Much basic research has been done to improve the man-machine interface, and a good deal of this research has employed classification and analysis of human errors as one element of the effort, often a significant element. This thesis draws heavily on the extensive work done by William B. Rouse and his associates, Sandra H. Rouse, Joost M. Van Eekhout, William B. Johnson, and Jens Rasmussen, Rouse and Rouse [83], Van Eekhout and Rouse [81], Johnson and Rouse [82], Rasmussen and Rouse [81]. The material in the first section of this chapter relies on their work in the cited references and is the basis for Chapter Three of this study. Chapter Three will develop the error classification scheme for this study. However, first it is important to understand the development of a process model and error classification scheme in the general sense.

# CLASSIFICATION OF HUMAN ERROR

## Human Errors

Human errors can be described and categorized in a number of ways. This discussion is limited to the ways that appear to be most relevant to the topic of this thesis, mechanic's errors. If the premise is accepted that errors are the result of causes, and not simply random occurences due to chance, then the analysis of errors can take the path of linking the errors and the causes.

Error analysis can be used to examine the worth of a management system of a training program or of whatever structure is to be evaluated. This can be done regardless of whether the structure in question is to be evaluated on its own merit or compared to another concept for the same structure. The error analysis here will examine the Army system on its own merit. The effort will be to tie errors to causes, through contributing factors.

As was noted in the first chapter, if the interest is the evaluation of a system such as that within which mechanics are selected, trained, employed, and provisioned, then the observation must be of mechanics who are products of that system and who are working in that system at the time they are observed.

## The Process Model

As with most technical skills, there is a procedure followed by a vehicle mechanic in accomplishing a diagnostic and repair task. This procedure involves separable steps that require the mechanic to observe, make choices, and act. Taking these steps collectively, the mechanic applies skill and knowledge. If errors are to be categorized in a meaningful manner, the error should be identified as occuring within one of these procedural steps. In that way, the error can be analyzed in light of what the mechanic was supposed to have done at that step. This provides the opportunity to identify error categories within the error classification scheme, and assists in identifying relationships within categories.

Rouse depicts the model for an aircraft pilot in Figure 2-1. This model can be restructured to fit the case of the mechanic. For the mechanic, the first decision is to determine if the vehicle works properly, and if not, what the symptoms are. Based on this, the mechanic will decide what test to use, what to fix based on the test result, and finally, how to fix it.

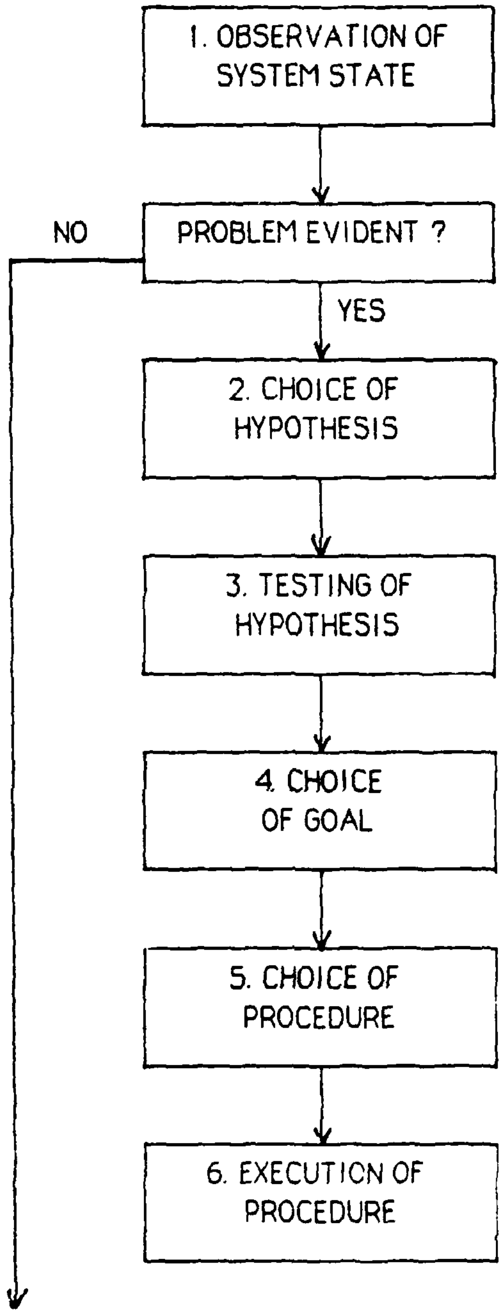



Figure 2-1

Conceptual Model of Pilot's Tasks, Rouse and Rouse [83]

It is easy to see that there are errors that would be associated with each of these three steps.

## The Classification Scheme

Table 2-1 shows a portion of Rouse's error classification scheme for Figure 2-1. There are two categories of errors, general and specific. The general category sorts the behavioral processes that would be associated with the error, and the specific category defines the characteristic of a wrong human action or decision.

An example of how an error is categorized is as follows: an observer watches a mechanic performing a vehicle check to determine if symptoms exist. An error occurs if the mechanic makes an incomplete observation of the symptom. If the mechanic sees that one brake is dragging, does he check to see if the other brakes are dragging? These two symptoms suggest different hypotheses. The general category of error is the behavioral process of a wrong observation, the specific category error is that the observation is incomplete.

An error classification scheme should include as many specific category listings as is necessary to describe these wrong behavioral actions. Completeness here will greatly influence the quality of the post-test analysis of errors types. Table 2-2 lists brief definitions of the specific categories from Table 2-1.

Table 2-1

Error Classification Scheme, Rouse and Rouse [83]

| General Category | Specific Category |
| --- | --- |
| 3. Test Hypothesis | a. incomplete<br>b. false acceptance of the wrong hypothsis<br>c. false rejection of the right hypothesis<br>d. lack |
| 4. Choose goal | a. incomplete<br>b. incorrect<br>c. unnecessary<br>d. lack |
| 5. Choose Procedure | a. incomplete<br>b. incorrect<br>c. unnecessary<br>d. lack |

## Causes and Contributing Factors

Causes for errors are suggested by the final analysis of the test data. However, as will be seen in the section on design and analysis of the experiment, an idea of contributing factors to be considered must be present during the test design-phase. Consequently, some consideration must be

Table 2-2

Definitions of Specific Error Categories, Rouse and Rouse [83]

| Specific Category | Brief Definition |
|---|---|
| 3a. incomplete | stopped before reaching a conclusion |
| 3b. acceptance | reached wrong conclusion |
| 3c. rejection | considered, discarded right conclusion |
| 3d. lack | hypothesis not tested |
| 4a. incomplete | insufficient specification of goal |
| 4b. incorrect | choice of counter-productive goal |
| 4c. unnecessary | choice of non-productive goal |
| 4d. lack | goal not chosen |
| 5a. incomplete | choice would not fully acheive goal |
| 5b. incorrect | choice would acheive incorrect goal |
| 5c. unnecessary | choice unnecessary for acheiving goal |

given to error causes in defining both the hypothesis and the type of test to conduct.

The inclusion of specific as well as general categories of errors can aid in the development of post-test hypotheses for further study. Every classification should be considered on its own merit. However, no classification should be discarded because there is no obvious causal relationship prior to the test.

It will also serve the purposes of the test to carefully develop the list of contributing factors to be considered by the observers during the conduct of the test. Factors such as basic knowledge, distraction, workload stress, and environment should all be considered.

It is important to gain a consensus of experts regarding issues of error categories, contributing factors, and causal factors. This will reinforce the value of the test result, as power is added to the result if a consensus of experts supports the rationale going into the test.

## Post-Test Analysis of Errors

For the sake of test validity, the previously mentioned experts and observers must be two separate groups of people. The observers must be experts in vehicle maintenance, but they must also be trained to record information with as little tendency for bias as is possible. The observers must be trained to be as neutral as possible, so that their observations are not influenced in either a positive or a negative fashion. If this is not the case, the test result may be so biased as to be of no value, or worse yet, it may result in undetected erroneous conclusions that would result in the loss of more money than did the original problem, had it been left alone.

The group of experts who validated the error categories and contributing factors should also do the post-test error scoring. They should use the observer's reports to do this. The first step in the error-validation process is to identify an error and to conclude that it will be considered rather than discarded, as its being a matter of preference difference on the part of the mechanic, or as its being irrelevant to the process model. The next step is to score the remaining errors by judging which contributing factor is most responsible. Disagreements must again be resolved by consensus. This is done after each expert has scored all of the errors independently. Statistical analysis can only begin after the error scoring phase is complete.

## DESIGN AND ANALYSIS OF AN EXPERIMENT

### The Benefit of an Experimental Design

The ability to arrive at a conclusion based on the statistical analysis of data is a function of how varied that data is in its reflection of the reality that it is supposed to represent. Like a distorted mirror, data can present such a poor image as to be useless in discovering anything about the case in point. Continuing the same analogy of the mirror, a similar problem, often coincidental, is the strength of light in the room. If the light is poor, even a good mirror presents an indistinguishable image.

The first part of the analogy with the mirror refers to noise, the second refers to volume. The image that the data presents can be enhanced by reducing the distortion (noise), and by turning up the light (volume). In principle this is a simple idea, in fact it is sometimes difficult, but nonetheless is a real possibility. The key to this possibility is experimental design.

The development of controlled experiments can hold at bay much of the potential noise while simultaneously working to turn up the volume. Some experimental designs work primarily on noise reduction, others turn up the volume, some do both. The key is to select a design that is both appropriate to the hypothesis being examined, and one that gives the greatest return in the strength of the potential conclusions. Chapter Four will present the experimental design for this study and the recommended means for the analysis.

## The Hypothesis

An experiment must be structured around a hypothesis. Statistics relies on a hypothesis that defines a situation, and then attempts to determine, with some stated degree of certainty, whether or not the hypothesis, as stated, is true or false. In the case of determining whether or not the errors that mechanics make are the result of controllable factors, it would be appropriate to state the hypothesis as follows: the errors that mechanics make are random and are not significantly correlated with any independent (contributing) factors.

The direction that the test design must take is that of formulating a scenario for data collection that will allow conclusions to be drawn on whether or not the hypothesis is correct and with what degree of certainty. The previous discussion of error categories, error frequencies, contributing factors, and causes, can all be addressed by addressing the stated hypothesis with a designed experiment.

## Reducing Uncontrolled Factors

The reduction of uncontrolled factors will be one of the tougher elements to deal with in the design of the experiment. In the effort to observe mechanics in the working environment, and to be able to draw conclusions about unsatisfactory influences in the system, influences not to be considered must be controlled. As an example, to determine the value of one type of training over another, the subjects of the test should have all things in common except the type of training. In reality, this is not totally possible, nor totally necessary, but all potential differences and their impacts must be examined. Those that are judged to be not significant are ignored. Those that are judged to be significant must be controlled by the test, or adjusted for after the test. The difficulty is in identifying these factors for, once identified, it is usually possible to negate their effects.

## Experiment Design

The experiment design is not arrived at in isolation of the hypothesis that is being examined. An experimental null hypothesis which has been carefully decided, dictates the experiment design.

In this thesis there is a primary experimental null hypothesis that will be addressed by the collection, reduction, and subsequent analysis of data. The data will be collected by trained people working in a predetermined sample environment. The control over the experiment is intended to insure that the results are valid, and to insure, if the experiment is run again, that the experiment result is confirmed.

The secondary hypothesis has to do with treatments. In Chapter One, it was mentioned that it may be of interest to know whether the frequency of mechanic's errors increased when the mechanic moved from garrison to the field. And if so, were the correlations the same or different. To examine this question, the entire primary experiment must be done over but now in a field environment, assuming the first experiment was done in garrison. The treatment difference (garrison versus field) should reinforce the correlation result of the first experiment, by replication, and go on to answer the questions regarding increased error frequencies and new or different correlations.

## Statistical Analysis

There are three types of statistical analysis suggested in the work done by Rouse. They are the three recommended by the author of this thesis to evaluate the data in different ways. A paired *t*-test is used to examine between-group differences where there are two treatments. Correlation methods are used to assess the relation between error frequencies and contributing factors. And finally, if the frequencies of errors are found to be correlated, multivariate analysis of variance is used to analyze the multiple dependent measures provided by the frequencies in the different error categories.

The decision regarding which statistical method to use is dependent on the experiment design used. The tools of inference-making, estimation, and confidence bounds, must all be examined for an appropriate place in the analysis of the test result.

# Chapter 3

# ERROR TYPES AND CONTRIBUTING FACTORS

In this chapter the process model example of Chapter Two is used as the basis for the development of a vehicle repair process model. This process model is then used to identify the verbs associated with the process steps. These verbs are the basis for the general error categories. Modifying words are selected which describe erroneous human actions. In the second portion of the chapter, contributing factors are examined to determine the influences that impinge on a mechanic's performance.

## THE PROCESS MODEL

What are the required actions or steps that accompany a complete diagnosis and restoration of a malfunctioning vehicle? The mechanic is ordinarily told of symptoms or likely faults or both. At this point the mechanic will employ knowledge, judgement, and experience acquired in training and on the job, in the use of his senses, test equipment, and manuals to diagnose and restore the vehicle to an operational state. This process can be described by a set of discrete events which are successive, involve decisions, and provide feedback.

Development of this process model underlies the preparation of the error classification scheme. The steps or events in the process model become the general error categories. The process model is shown at Figure 3-1.

Observe the Equipment

The first step in the process, in the case of unscheduled service (viz. a vehicle breakdown), is to observe the operation of the vehicle in light of any information provided by the vehicle's operator or the mechanic's supervisor. In the case of scheduled service (viz. periodic maintenance), symptoms will be made apparent by the prescribed vehicle check-out procedure. In the case of either scheduled or unscheduled service, the mechanic must determine whether adequate symptoms of a vehicle malfunction exist or do not exist. The verb of interest in this step is "observe". This step will be followed by a decision which is based on the quality of the observation. It should be noted that a correct decision on the

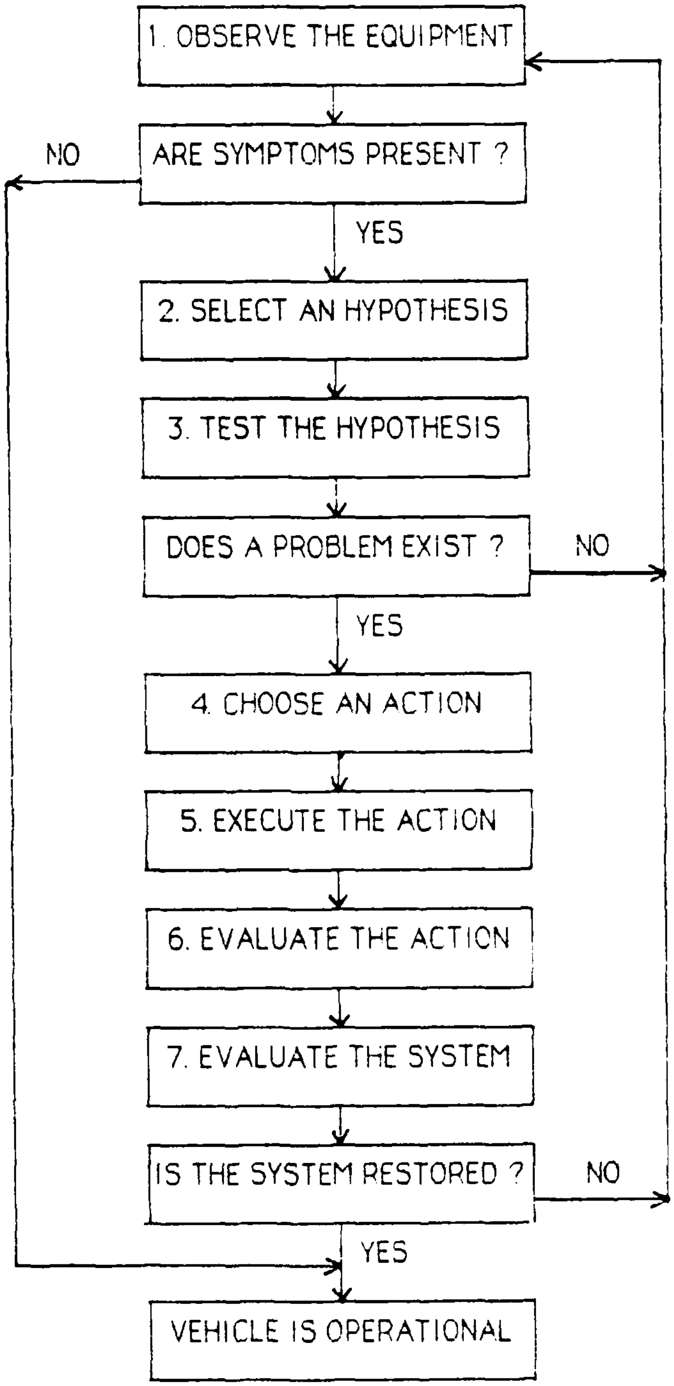



Figure 3-1

The Repair Process Model

part of the mechanic could be that no symptoms exist, and the consequence is that the entire process is halted. This possibility is indicated by the process bypass at the decision block following the first step.

## Select the Hypothesis

Once the symptoms are observed, it is necessary to determine whether the system is malfunctioning. This may not be trivial due to the interconnection and dependency of systems. An example of such a situation is of an engine stalling. This could be due to the ignition sub-system, the fuel sub-system, the engine mechanical sub-system, or some combination of these sub-systems. The mechanic must select a reasonable hypothesis based on the symptoms that are observed. This selection process is often iterative and is based on test-result feedback. The fact that hypotheses are selected, tested, and rejected iteratively is part of the diagnostic portion of the total process. The wrong selection of an hypothesis is not a selection error in itself, unless it was selected on erroneous grounds. The verb "select" is used to define this step.

## Test the Hypothesis

Testing a selected hypothesis can be as simple as looking, listening, touching, or smelling. On the other hand, testing may take on larger proportions, by including the application of tools, techniques, and testing and measuring equipment. Hypothesis testing may involve the execution of a considerable number of steps which include selecting actions, making decisions, and drawing conclusions. In the end, the test is followed by a decision as to whether or not the hypothesized problem exists. As mentioned previously, the hypothesis selection step and the hypothesis testing step are often iterative. This is a normal part of the total process and is described with a feedback loop. In this step the verb "test" is the word of interest.

## Select an Action

After the determination has been made that a problem exists, an appropriate action must be selected that will correct the sub-system malfunction. This action may involve a repair, a replacement, an adjustment, or a combination of these actions. Regardless of the action chosen, the verb that is central to this step is "select". Accomplishing these actions is often a process in itself and will be covered in the next step.

## Execute the Action

Executing an action that restores the sub-system function may involve the expenditure of considerable time and the application of a great deal of skill. Complex actions in this step are often the root-cause of additional sub-system malfunctions. Just as unnecessary repairs waste time and parts, incorrect repairs result in the same loss of resources, but often go further and result in immediate or future failures that may range in nature from minor to catastrophic. The mechanic must employ considerable skill, judgement, and experience in this step, just as he did in the second step, that of selecting the hypothesis.

Evaluate the Action

A necessary step that is often interleaved with the preceding step of action taking, is for the mechanic to check his work as he proceeds through system restoration. It is often not possible to accomplish this step after the repair action is completed. This step should not be confused with the next step of system evaluation. Action evaluation involves checks as does the next step. However, these steps are usually separate and distinct process-steps. Checking the tightness of a bolt or a nut during re-assembly is part of this step; testing for brake sub-system functioning is not part of this step. "Evaluate" is the verb of interest.

## Evaluate the System

As was mentioned, the evaluation of the system function is a separate step apart from evaluation of the repair action. Very often, the system evaluation uses a prescribed check-out procedure which the mechanic employs to verify not only that his repair did not result in new problems, but also to insure that no additional or related symptoms persist. It could be that the symptoms of the corrected malfunction masked the symptoms of other malfunctions. In this case the mechanic may find that he has additional symptoms that will provide feedback for re-execution of the process starting at step one. This decision to return to the first step is based on the findings of the system evaluation and is defined by the feedback loop to step one. As in the preceding step the verb of interest is "evaluate".

The above concludes the discussion of the process model with the exception of a final note. In the workplace, steps are often seen to overlap or interleave. It is important that this phenomenon is not seen as representative of erroneous action inasmuch as this is a natural process that does not embody the nature of the discrete event that is suggested by the step-by-step process model illustrated.

## THE ERROR CLASSIFICATION SCHEME

Now that the process model is fully described, the general error categories are simply taken from the word phrase that contains the verb of interest in each process step. The decision nodes are not ignored. Rather, an erroneous decision is a specific category error in the general category that precedes the decision node. It is now necessary to develop a list of words that can be used to modify the selected verbs or their noun forms to describe a related erroneous human action. This list should be developed for each general category verb. Table 3-1 shows the adverbs and adjectives that are associated with the verbs observe, select, test, execute, and evaluate. These adjectives and adverbs are then used to develop a list of erroneous actions that can be related to the activity described by the general category error descriptions. This list is shown as Table 3-2.

Table 3-1

Error Classification Scheme

| General Category | Specific Category |
|---|---|
| 1. Observe the Equipment | a. misinterpreted<br>b. incorrect<br>c. incomplete<br>d. inappropriate<br>e. missing |
| 2. Select the Hypothesis | a. inconsistent<br>b. unlikely<br>c. costly<br>d. irrelevant<br>e. missing<br>f. incomplete |
| 3. Test the Hypothesis | a. incomplete<br>b. incorrect acceptance of wrong hypothesis<br>c. incorrect rejection of right hypothesis<br>d. missing<br>e. irrelevant<br>f. wrong |
| 4. Select an Action | a. incomplete<br>b. incorrect<br>c. irrelevant<br>d. missing<br>e. unlikely<br>f. excessive |

Table 3-1 Continued

Error Classification Scheme

| General Category | Specific Category |
|---|---|
| 5. Execute the Action | a. step omitted<br>b. step repeated<br>c. step added<br>d. incorrect<br>e. inappropriate timing<br>f. incomplete step<br>g. unrelated<br>h. incomplete sequence |
| 6. Evaluate the Action | a. incomplete<br>b. missing<br>c. incorrect<br>d. unnecessary<br>e. misinterpreted |
| 7. Evaluate the System | a. misinterpreted<br>b. incorrect<br>c. incomplete<br>d. unnecessary<br>e. missing |

Table 3-2

Definitions of Specific Error Categories

| Specific | Description |
|---|---|
| 1a. misinterpreted | observed an occurring symptom but misinterpreted it |
| 1b. incorrect | identified a symptom that was not observed |
| 1c. incomplete | did not observe all occurring symptoms |
| 1d. inappropriate | used inappropriate means to observe a symptom |
| 1e. missing | failed to observe an occurring symptom |
| 2a. inconsistent | symptoms do not suggest this hypothesis |
| 2b. unlikely | more likely hypotheses based on existing symptoms are available |
| 2c. costly | hypotheses which are less time consuming to test are consistent with these symptoms |
| 2d. irrelevant | hypothesis not functionally related to symptoms |
| 2e. missing | failed to select an hypothesis |
| 2f. incomplete | hypothesis does not account for all symptoms |
| 3a. incomplete | stopped prior to reaching a conclusion |
| 3b. incorrect acceptance of wrong hypothesis | reached the wrong conclusion |
| 3c. incorrect rejection of right hypothesis | discarded the right conclusion |
| 3d. missing | did no test |
| 3e. irrelevant | tested a hypothesis other than the one chosen |
| 3f. wrong | correct test done wrong |

Table 3-2 Continued

Definitions of Specific Error Categories

| Specific | Description |
|---|---|
| 4a. incomplete | chosen action can only partially fix identified problem |
| 4b. incorrect | chosen action will make problem worse |
| 4c. irrelevant | chosen action irrelevant to identified problem |
| 4d. missing | failed to choose an action |
| 4e. unlikely | chosen action unlikely to fix identified problem |
| 4f. excessive | chosen action goes beyond what is required |
| 5a. step omitted | required step left out |
| 5b. step repeated | required step repeated unnecessarily |
| 5c. step added | unnecessary step added |
| 5d. incorrect | step executed incorrectly |
| 5e. inappropriate timing | required step completed out of sequence |
| 5f. incomplete step | did not finish a required step |
| 5g. unrelated | unrelated action executed during sequence |
| 5h. incomplete sequence | stopped prior to sequence completion |
| 6a. incomplete | did not fully evaluate completion of action |
| 6b. missing | did not evaluate completion of action |
| 6c. incorrect | incorrect evaluation of action |
| 6d. unnecessary | evaluated unrelated action |
| 6d. misinterpreted | misinterpreted evaluation result |

Table 3-2 Continued

Definitions of Specific Error Categories

| Specific | Description |
|---|---|
| 7a. missing | did not evaluate system function |
| 7b. incorrect | incorrect evaluation of system function |
| 7c. unnecessary | evaluated unrelated system |
| 7d. misinterpreted | misinterpreted evaluation result |
| 7e. incomplete | did not fully evaluate system function |

## CONTRIBUTING FACTORS

Careful development of the list of contributing factors is of great importance to the success of the proposed experiment. The factors that are seen as contributing to the observed errors provide the link to the cause of the error. Unless the errors and their frequencies are related to a systemic weakness described by a contributing factor, it must be concluded that the errors are random and not significantly correlated with any contributing factors. For this experiment, and for the subsequent analysis to provide a substantive result, the factors that contribute to the errors must be tied directly to some human or mechanical technology that has the potential for improvement.

The phrase, human or mechanical technology relates to the various technological activities that influence the selection, training, employment, and provisioning of the soldier mechanic. These include psychological testing which leads to the identification of mechanical aptitude; institutional training which includes such technologies as design of instructional programs, design of instructional materials, and computer aided instruction; maintenance management which includes management of human and material resources; equipment repairability or maintainability which are technologies in themselves; and finally, the technology involved in the design of test equipment to aid in the diagnosis and repair of electrical, hydraulic, and mechanical systems.

Taking these groupings one at a time, it will be necessary to consider exactly how the contributing factors may exhibit themselves. In addition to selection, training, employment, and provisioning, will be added the environmental and physical/mental factors.

Table 3-3

Contributing Factors

| General | Specific |
|---|---|
| Selection | 1. Inadequate Mental Ability<br>2. Inadequate Physical Ability |
| Training | 3. Inadequate Initial Entry Training<br>4. Inadequate Unit Training |
| Employment | 5. Incorrect Employment<br>6. Inadequate Supervision |
| Provisioning | 7. Inadequate Manual<br>8. Inadequate TMDE*<br>9. Inadequate Tools<br>10. Incorrect Use of Manuals<br>11. Incorrect Use of TMDE*<br>12. Incorrect Use of Tools<br>13. Missing or Incomplete Manual<br>14. Missing or Inoperative TMDE*<br>15. Missing or Broken Tools |
| Personal Factors | 16. Fatigue<br>17. Stress<br>18. Distraction<br>19. Tension |
| Environmental Factors | 20. Weather, Field Environment, Workplace Inadequacy, Workplace Confusion |
| Vehicle | 21. Vehicle Repairability |

*TMDE. Test Measurement and Diagnostic Equipment

### Selection

Selection of an individual to be a soldier mechanic is based upon the mental and physical attributes of the individual applicant in relation to what physical and mental attributes a soldier mechanic should possess. The potential selectee is administered the Armed Services Vocational Aptitude Battery of tests to determine mental aptitudes, and is given a general physical examination to detect physical limitations as well as physical strength. The qualifying attribute scores are established by the US Army. These qualifying scores are based upon research which has established what humans are capable of accomplishing, given specific physical and mental assets, coupled with research into what a soldier mechanic must be capable of doing to accomplish his mission. Thus, there are two technologies involved, psychology and medicine. The contributing factors that relate these technologies to an error on the part of the mechanic are inadequate physical ability and inadequate mental ability. It can be assumed that since the individual being observed is a mechanic in the US Army , he possesses the required mental and physical attributes for the position. Therefore errors attributed to one of these factors suggests that either the evaluation is inadequate or the requirement is inadequate. These factors can easily overlap with physical fatigue or mental stress which could be due to other factors, is examined also.

## *Training*

Training, and the support of training, encompasses a large number of interrelated fields of technology. They include the design of training programs, design of training materials, computer-assisted instruction, and the testing for comprehension and retention. Supporting these disciplines are psychology, computer science, electronics, mechanical and electrical engineering, and other technologies. From the perspective of training, or the lack thereof, being a contributing factor to the mechanic's making errors, only his supposed knowledge base is of interest. The soldier mechanic has either comprehended and retained the information that he was exposed to in the institutional environment or he has not. Consequently, to a trained observer, a judgement will be required as to whether or not an error was related to a lack of knowledge that should have been gained in training. If the observer is well informed of the skills that are taught in the institutional environment and those that are taught in the unit environment it may be possible to separate the two. However, if the mechanic is far removed from his institutional training in time, such as two or more years, it will be difficult to fault the institutional training. The observer must be aware of how long it has been since the mechanic received Initial Entry

Training and what this training consisted of. It will be appropriate for the observer to list institutional or unit training as a contributing factor and note any special circumstances which he, the observer, is aware of that may aid the panel of experts in further defining the error.

## Employment

The area of employment of the mechanic presumes that he is assigned to a position that requires his particular Military Occupational Specialty. Beyond this factor is the proper task assignment commensurate with his skill level and proper supervision by an individual who is qualified to supervise the mechanic. In this case, the contributing factor would be incorrect employment or inadequate supervision. An entry level mechanic, Paygrade E1 to E4, must have some access to his supervisor to confer on problems related to his assigned tasks. This access applies regardless of whether the mechanic is working in garrison or in a field location. Here again, the observer must be familiar with the applicable Soldier's Manual to be aware of what the mechanic is expected to accomplish.

## Provisioning

This area covers three separate types of factors that can be distributed over three sorts of material provisioning. The three types of factors are: one, inadequacy of the supplied item, two, incorrect use of the supplied item, and three, lack of the required item. The three item-groups are: one, Technical Manuals and other printed aids; two, test, measurement, and diagnostic equipment; and three, ordinary and special tools.

The first group includes troubleshooting, maintenance, and parts manuals; lubrication orders; technical bulletins; and other printed aids. Test, measuring, and diagnostic equipment covers all mechanical, hydraulic, and electrical devices which require maintenance and/or calibration. Items such as torque wrenches, meters, and automatic test equipment fall into this category. The third group encompasses the tools that are supposed to be in the mechanic's personal tool set and the tools that are supposed to be in the shop's tool sets. In each case the observer must determine if the correct tool is available, whether or not it is used correctly, and finally, whether it is adequate relative to what it is supposed to accomplish.

### Personal Factors

The observer must be able to comment on the apparent state of the mechanic in terms of his capability to function normally. This observation should be accurate enough to evaluate the extent of such factors as fatigue, stress, tension, and distraction as contributors to error. Although these factors may be closely related to other factors, or a result of other factors, the observer must determine the leading contributing factor in each case of a suspected error incident.

Environmental Factors

This final area, environmental factors, is limited to one category although it may encompass several factors. A mechanic must be able to function in a variety of severe and hostile environments. However, to complete the list of contributing factors, the observer must be able to account for weather, work-place confusion, or work-place inadequacy, hostile field environment, and other such related factors. The observer must again use judgement when identifying environmental factors as the leading contributor to an erroneous action.

## CONCLUSION

The various specific factors contributing to mechanics errors may be aggregated after the experiment to improve correlation at a more general level. However, it is important to be as discriminating as possible during the conduct of the experiment to isolate specific contributing factors for which there is a priori reason to believe that correlation exists. This will be discussed further in Chapter Four under correlation.

Table 3-4

Vehicle Subsystems

| Group | Sub-System |
|---|---|
| Drive Train | 1. Engine<br>2. Fuel<br>3. Cooling<br>4. Clutch<br>5. Transmission<br>6. Transfer<br>7. Propeller Shafts |
| Other Mechanical, Electrical, or Hydraulic | 8. Axles<br>9. Brakes<br>10. Wheels<br>11. Steering<br>12. Suspension<br>13. Electrical |
| Other | 14. Exhaust<br>15. Frame, Towing, Winch<br>16. Body, Cab, Hood<br>17. Non-Electric Gauges<br>18. Special |

# Chapter 4

## DESIGN AND ANALYSIS OF THE EXPERIMENT

An experimental design is used to obtain analytically useful information in an economically prudent manner. The statistical analysis techniques used should obtain as significant a result as is possible. At the same time, this effort should be accomplished with the least amount of information as possible. It is the sample data that is being purchased in an experiment and, thus, the smaller the sample the less the cost.

The first decision regarding the design of an experiment is the development of the hypothesis. Next is the determination of the appropriate methods of statistical analysis to employ in testing that hypothesis. The chosen statistical method will drive the decision on what kind and how much information is necessary for arriving at a statistically significant conclusion. The experimental design specifies how the information will be collected, under what circumstances it will be collected, and how much of it to collect. The details of the experimental design ensure that the information collected accurately represents the total population of interest.

Randomness of the selection of test subjects protects the information from being non-representative, while tests on the information collected rechecks this validity.

Each step of this process will now be discussed in the order just presented. The conclusion of this chapter will touch on other considerations that fall just outside of this present discussion.

## THE HYPOTHESIS

The hypothesis addresses the question of whether or not there is a relationship between the independent contributing factors and the dependent types of errors committed by mechanics. If this relationship exists and the connection between contributing factors and associated technologies exists, a correct statistical analysis will reveal this relationship as well as the relative contributions of the technological weaknesses to the errors made by mechanics, Van Eekhout and Rouse [81].

The accuracy of the statement of the hypothesis underlies the development of the experimental design and directs the statistical analysis. In this study, the working hypothesis is that mechanics make errors, these errors stem from the presence of contributing factors rather than being random errors, and that these contributing factors relate to weaknesses in the system that selects, trains, employs, and provisions the mechanic. This working hypothesis will form the basis for the experimental null and alternate hypotheses. The null hypothesis, the hypothesis to be disproven if the working hypothesis is true, states that mechanic's errors are random. The alternate hypothesis is a formal statement of the working hypothesis.

## The Null Hypothesis

If a soldier meets the existing mechanic's selection criteria, successfully completes the prescribed mechanic's training, is employed and supervised in his assigned unit in accordance with applicable US Army policies, and is supplied with the required tools, test equipment and manuals, the only errors that he will make in the diagnosis and repair of vehicles that he is responsible for will be random errors which are not significantly correlated with any contributing factors.

## The Alternate Hypothesis

In spite of the proper use of the current selection, training, employment, and provisioning practices, errors committed by mechanics are not random. Further, these errors are correlated with contributing factors that represent weaknesses in the system that selects, trains, employs, and provisions soldier mechanics.

## EXPERIMENTAL ERROR

The leading danger in any test of an experimental hypothesis is that the wrong conclusion will be drawn. Acceptance of the alternate hypothesis when in reality the null hypothesis is the true state of nature is called a Type I error. There is a recognized effect that occurs when an observer watches an individual perform a task and the observed individual is aware of the observer's presence, this effect is called the "Hawthorne Effect". The task performer, in this case the mechanic, will tend on the average to do better work when observed than when not observed regardless of whether he resents or enjoys the attention. In this experiment, this effect will reduce error frequency and thus tend to decrease the likelihood of a Type I error. Of course, the converse of this situation, acceptance of the null hypothesis when the alternative hypothesis is correct, a Type II error, is a danger if the observers are overly strict and assess too many errors. If a great number of trivial errors are included, a conclusion may be reached that establishes a level of correlation that does not exist. To avoid this danger, the observer's judgement must not be used in determining whether or not errors have occurred. A separate panel of experts will independently evaluate and determine the error cases and associated contributing factors. These same experts must then reach a consensus regarding all errors to be considered in the analysis. This protection will lead the analysis away from the Type II error. However, in this experiment the consequence of a Type II error is preferred over that of a Type I error. The reason for this is that acceptance of the alternate hypothesis will result in a change from the status quo. It should be obvious that the current situation is preferred to a change, when the change would be an error. Therefor, slight bias in favor of the consequence of a Type II error is preferred.

## THE REQUIRED ANALYSIS

This experiment described in this thesis is directed towards discovery of a relationship between types of errors which are effects, and contributing factors which are believed to be the causes. However, the errors could be totally random in nature with no clear-cut relationship between the errors and the contributing factors.

In this experiment, the error classification scheme seeks to break down all of the errors made by mechanics into specific types to better address the causes. This classification results in the sum of all errors being subdivided into a set of discrete elements which are thought to be dependent on a similarly constructed set of discrete contributing factors, which are the independent factors. If the relationship exists, it can be described with the typical linear regression model:

$$Y = \beta_0 + \beta x + \epsilon$$

where:

$Y$ = the total number of errors made by the mechanic

$\beta_0$ = the average number of errors due to random causes

$\beta$ = the number of errors due to contributing factors

$x$ = 1 if contributing factors exist, 0 otherwise

$\epsilon$ = error term

for canonical correlation:

$$Y = \alpha_1 Y_1 + \alpha_2 Y_2 + \ldots + \alpha_n Y_n$$

$Y_j$ = error type j  j=1,2, . . . ,n

$Y_j$ = 1 if error type j exists, 0 otherwise

$$X = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m$$

$X_i$ = contributing factor i  i=1,2, . . . ,m

$X_i$ = 1 if contributing factor i exists, 0 otherwise

Furthermore, the analysis must provide a means to relate the influences among the $X_i$'s on the set of Y's. This type of analysis can only be accomplished with canonical correlation.

Canonical correlation was developed by Hotelling in the mid 1930's, Lindeman [80]. The first item of interest to the experimental design relative to the use of canonical correlation is the sample size. If there are $p$ contributing factors and $q$ error types to be worked upon, the sample size $n$, may be determined as Lindeman[80]:

$$n = 20(p+q)$$

Chapter Three identified 41 error types and 21 contributing factors. This would lead to a minimum sample size of:

$$n = 20(41+21) = 1240 \text{ errors}$$

Given that mechanics commit an average of two errors per repair effort, Kern [UN], and that mechanics perform six to seven repair tasks per day (an estimate based upon personal experience), the required sample size is equivalent to approximately 100 mechanic-days of observation. This sample size could be obtained by observing ten mechanics for ten working days, which is two calendar weeks.

Of consideration at this point is how long an observer may spend with a given mechanic prior to the observer losing his objectivity. If the observer tends to become sympathetic to the mechanic as the observation period grows, it would be necessary to argue in favor of limiting the period to only one week per mechanic. This would require the observation of twenty mechanics. Depending on the observer resources available, the total period for observation could range from one to several weeks.

A larger randomly-selected sample will provide a better representation of its respective total population, arguing in favor of a greater rather than lesser number of mechanics. This being the case, twenty mechanics for one week would be preferable to ten mechanics for two weeks.

The error cases collected will be subjected to screening on two levels. The first level was discussed in Chapter Three. Rouse indicated that virtually all identified errors were retained for the sample due to the ability of the expert judges to reach agreement on all cases, Rouse and Rouse [83]. Kern had the observers list their estimates of error occurences on their observation records, and then the records were submitted for judging, Kern [83]. Although no mention was made of errors lost due to judging, the previously mentioned two errors per case was the number after judging. It can then be concluded that a sample of 1300 errors in 100 mechanic-days of observation should produce the required sample of 1240 errors after some judging losses.

There is another potential for loss of error cases and this will be discussed in the following section. Although no specific references were found in the literature, it is quite possible that one or more mechanics with their respective error sample could be removed due to their failure to meet the homogeneity test. As will be shown, a mechanic may be found to be non-representative of his population group due to his performance being too good or too bad. For this reason judgement dictates the addition of one or more extra mechanics to the sample for insurance against a too small final sample.

## HOMOGENEITY OF THE SAMPLE

In practice, randomness means that the sampling units (mechanics) of the observation vectors (set of error frequencies) were drawn independently of one another from some homogeneous population. With some limits on the range of the sample's randomness, there exists the likelihood that some sample data may fall outside those limits. Given that this experiment will include a randomly-selected group of mechanics, and that as such they should represent a larger group, it is possible that one or more of the sample group could perform so well or so poorly as to represent the fringe of the total population. In the sample of only twenty or more mechanics this fringe dweller could affect the desire for the sample to reflect the whole.

A test for accepting or rejecting a mechanic in the case of the number of errors committed in total would look at how many errors he committed relative to the average number of errors committed by the group. In this way a decision could be made to reject this mechanic due to his total error count being non-representative.

It can be decided on probabilistic grounds to reject any mechanic whose performance falls outside a stated bound. It could be decided that this bound would reject any mechanic whose performance places him in a small upper or lower fraction of all mechanics. The analysis should include only those mechanics who fall in the middle 95% of their population. This can be done using an F-test on the average total number of errors committed by the total sample of mechanics versus the total number committed by an individual mechanic, by applying the following hypothesis:

$H_0$: $\mu_1 = \mu_2 = \ldots = \mu_n$

If $H_0$ is rejected, remove the mechanic's error sample that deviates the furthest from the mean total number of errors and rerun the test. Do this until $H_0$ is accepted. Judgement must be used in selecting the acceptance bound on the F statistic so that it is a reasonable bound and insures that the non-representative mechanics are rejected.

Another reason for rejection would be that the distribution of errors across the major error categories is non-representative. Here, a Chi-Squared test would be used, this test accumulates the deviations of a mechanic's error frequencies by error category from the average error frequencies for the group. Again, this test can be used with a predetermined bound that rejects non-representative mechanics due to their error distribution residing at the fringe of the population of all mechanics. As with the F-test, judgement must be used to set an appropriate bound.

## FACTOR REDUCTION

The sample size discussion indicated the need for a minimum of twenty sample errors per factor to be considered in the canonical correlation. The power of correlation increases as the sample errors per factor increases. This being the case, it may be worthwhile to reduce the number of factors to be considered prior to doing the correlation analysis. There are several approaches to be considered.

The first possibility is to simply reject those factors (error type or contributing factor) that have a relatively very small contribution to every other factor of the opposing set. If, for example, physical ability was cited only a small fraction of the time for every error type, it could be discarded along with its related set of errors prior to any correlation analysis. This is not recommended, inasmuch as the interaction complexities of this experiment are too great to be seen by inspection, and in the end the canonical correlation analysis will reject insignificant factors.

Another reduction could be acheived by grouping of factors under major headings. This could be done by using the seven major categories of errors and the seven major categories of contributing factors. But again, the complex interplay of the total number of error types and contributing factors would be lost. However, a statistical method does exist that will not only group the factors into major groups, but do so without losing sight of the complex interactions. This method is "factor analysis", and it will be covered in the following section.

## FACTOR ANALYSIS

The major purpose of factor analysis is to determine whether a set of variables can be described in terms of a number of super variables of less number than the set of all of the initial variables considered. The problem is to determine whether the $n$ variables in the larger set can be broken down into $m$ subsets. This aggregation of the $n$ variables is based on their exibiting a stronger relationship to the variables within their subset than to the variables not in their subset, Lindeman [80].

There are two types of factor analysis approaches. The first type is confirmatory factor analysis, the second is exploratory factor analysis. The first type assumes that the investigator has an idea of how the set of $n$ variables may be grouped into a hierarchy of subsets where the subsets express larger relationships than the more specific individual factors alone. Exploratory factor analysis makes the same assumption regarding subsets or factor groupings, but assumes no knowledge on which variables should be grouped together prior to the experiment. The second type of factor analysis applies here because the subsets of variables may cut across the major contributing factor categories. It is presumed, that if a relationship exists between error types and contributing factors, then certain groups of error types are related to certain groups of contributing factors.

To accomplish this statistical analysis it is necessary to produce a correlation matrix of all error types and a correlation matrix of all contributing factors. These matrices are operated on separately. If elements within a given matrix may be grouped with regard to correlation coefficients, then the grouped or correlated factors may reveal a meaningful relationship leading to additional hypotheses for test. This type of factor analysis is never an end in itself; rather it provides the potential for additional study topics from analysis of already collected data. Hence, the set of super factors would be used in the canonical correlation rather than the larger sets of all error types and all contributing factors.

Factor analysis can be summarized in the following five steps. First, a correlation matrix is developed. This matrix relates each variable in the set of $n$ variables to every other variable in that same set. Second, this matrix is worked on to determine a factor loading matrix. This matrix relates each variable to all others with weights which describe the relative strengths of correlation of the several variables to each other variable in the set.

In the third step the initial set of super factors is found using principal axis analysis or other similar method. This step can be related to regression analysis in that it seeks to find a set of principal axes or lines that represent as many of the $n$ variables and their relative relationships as possible. These lines or axes are orthogonal at first. The fourth step is to rotate these axes, to improve the quality of the expressed relationship between the principal axes and the subsets of the $n$ variables. This may be done retaining orthogonality or allowing the principal axes to define oblique angles. In either case, the rotation acheives the best fit of principal axes and the subsets of the $n$ variables.

The fifth step is for the investigator to study the newly created subsets and determine the characteristics that are being defined by each of the $m$ super factors. Inter-relationships between super factors are also quantified by factor analysis. These relationships must also be examined and explained.

Once the set of super factors is described, it must be determined if too much generalization has been acheived. If the final set of super factors for the larger set of $n$ contributing factors is too small, the connection between the contributing factors and their associated technologies may be lost. If this appears to be the situation with respect to contributing factors, the factor analysis effort should be abandoned in favor of going straight into the canonical correlation. Conversely, if error types are reduced to a small set of super factors the ability of the canonical correlation to acheive a meaningful result is not lost. Discrimination among the error types within the canonical correlation is not important to this analysis.

## CANONICAL CORRELATION

Canonical correlation will be used to determine if significant correlation exists between the several error types, the dependent variables, and the several contributing factors, the independent variables. This statistical method will go on to provide a vector of weights that represent the strengths of the correlations of the several contributing factors with the error types. The process of canonical correlation analysis is described in many texts covering multivariate analysis, the specific reference used here is Lindeman [80].

Canonical correlation produces a set of roots or eigenvalues, $\lambda_j$. This set contains a number of roots equal to or less than the number of variables in the smaller variable set used in the correlation. Only the first eigenvalue is of interest in this analysis. Additionally, because canonical correlation is an unrotated orthogonal projection, only the first eigenvalue is generally found to be significant.

The eigenvalue, $\lambda_1$, is used to develop a pair of canonical variates (eigenvectors), in this case $Z_{x1}$ and $Z_{y1}$. This pair of variates relates each

$X_i$ $(Z_{x1})$ to the set of Y's, and each $Y_j$ $(Z_{y1})$ to the set of X's. Both of these variates explains a percentage of the variance in the other variate. This percentage of explained variance is the eigenvalue from which the variates were derived. Each succeeding pair of variates $Z_{x2}$, $Z_{y2}$ and etc. , explains a percentage of the residual variance after removal of that variance explained by the preceeding pairs of variates. If the first eigenvalue is very high;

$$\lambda_1 \geq .90$$

for example, the remaining variate pairs are of little consequence. This is ordinarily the case if the significance of $\lambda_1$ exceeds the Wilk's Lambda test statistic at a significance of 0.05. Summarizing, if $\lambda_1$ is significant, it is often the only root that is significant. If it is not significant at the 0.05 level, there is no significant correlation in the system, and $H_0$ is not rejected.

The eigenvector or set of signed weights from variate $Z_{x1}$ provides a means to rank order the influences of each $X_i$ on the errors committed by the mechanics. The variate $Z_{y1}$ provides a means to rank the error types by those most influenced by the contributing factors. In this way it can be determined what proportion of the total errors fall in some upper portion of the ranked error types to determine what types of errors are most likely to be reduced if the effects of the highly ranked contributing factors are mitigated. The analytical results can be used to produce information for a decision breifing as shown in Appendix A.

## CAUSAL INFERENCE

Protection from wrong conclusions with regard to the relationships between error types and contributing factors is built into the canonical correlation method. However, there is no statistical test of whether the highly correlated contributing factors are related to specific weaknesses in the total system that selects, trains, employs, and provisions mechanics. This connection exists in the minds of the experts who ratified the contributing factors as being representative of weaknesses within the bigger system. Ultimately then, if the decision-maker or his representatives accept the development of the process model, the error classification scheme, and the contributing factors in Chapter Three, they should be able to accept the result of the canonical correlation.

## THE TYPE OF EXPERIMENT DESIGN

Canonical correlation can provide the analysis necessary to test the experimental hypothesis described in the first section of this chapter. In statistical analysis there is an opportunity to gain confidence in a result by repeating the experiment. In this case, the entire experiment can be run with two sample groups of mechanics, with each group of twenty or more mechanics being selected independently of the other and observed by different observer teams. These separate trials of the experiment can be done under different circumstances. Each experimental trial must stand on its own merit. If both trials reject the primary experimental null hypothesis, the subjective tendency is towards greater credibility, and the statistical fact is reinforcement of the significance of the findings. If both trials do not reject the primary experimental null hypothesis, an error has occurred and additional research must be accomplished to determine where that error lies.

As was said, the test of the hypothesis must hold up under any ordinary circumstance that finds military mechanics working on vehicles. The replication may be performed so as to not only test the same primary null hypothesis but also to examine a new factor. As an example, the replications could be conducted to examine wheeled-vehicle mechanic performance versus tracked-vehicle mechanic performance, or to examine the performance of mechanics working in garrison conditions versus field conditions. Many such comparisons could be envisioned. These replications would in fact be different treatments, and as such, would constitute an extension beyond the canonical correlation to examine a difference between treatments.

A comparison may be drawn between a garrison and field environment to determine if the correlations are similar as was suggested in Chapter One. If the two treatments, garrison and field environments, are examined it must be determined if there is total independence between the treatments or if there is some common link. The fact that the treatments would in most cases be variations in mechanic specialty or mechanic work environment, many factors are going to affect both treatments. Therefore, interaction between the treatments must be assumed and the design must be factorial. It stands to reason that if the experimental null hypothesis is rejected, the difference between treatments will reveal varying degrees in the effects of the contributing factors. The application of varied treatments will be discussed at the end of the chapter.

The experimental design employed to test the null hypothesis through the use of canonical correlation is not of the classic block or factorial designs. Rather, the experiment is designed around the hypothesis and the requirements of the statistical methods used. The classic regression model:

$$Y = \beta_0 + \beta x + \epsilon$$

would provide a naive model, as it could relate only one error type at a time to the contributing factors, or the total errors to the contributing factors. In both cases this approach would discount totally the simultaneous interaction of the independent contributing factors and the various types of errors.

## ADDITIONAL STATISTICS OF INTEREST

Earlier in this chapter, it was suggested that further worthwhile analyses were possible if consideration were given to the way in which the data were collected. It would be highly desirable to replicate this entire experiment to insure the validity of the result. Error classification schemes and the analysis of human error is a relatively new area with regard to rigorous statistical tests.

Because the experimental situation used here requires nc control group or other such device to insure statistical significance, it would be worthwhile to not simply replicate but to do so with a second treatment. Either trial should obtain a correct result on its own, and if the proposal used here is correct, the result should be comparable. A treatment difference could add another dimension to the result while satisfying the need to replicate. The recommended difference in treatment is to compare the garrison environment to the field environment. This would provide an interesting view of how the impact of the various contributing factors is intensified or relieved between the two environments.

Another statistical analysis of interest has to do with the vehicle sub-systems. The sub-system information collected by the observer can be used in companion analyses regarding the error frequencies by vehicle sub-system. This analysis could relate error types and vehicle sub-systems in much the same way that the principle analysis of this proposal did. Such an analysis could provide helpful information to mechanic trainers.

It would also be useful for the observers to collect the following item on their observation records. The observer could note whether or not a false replacement would have been the result of a recorded error. False replacement data may aid analysis when it is related to the information provided in Appendix A. The costs of false replacements could be indicated in Tables A-4 and A-5 to quantify the relationship between the contributing factors and the error types in terms of false replacement costs. Or, those errors that resulted in false replacements could be extracted from the total set of errors. This subset of errors that resulted in false replacements could then be used in a canonical correlation similar to the primary one.

# Chapter 5

## CONDUCTING THE EXPERIMENT

In this chapter the author discusses an approach to planning and executing the experiment. If there is a decision by the US Army to proceed on the basis of this thesis, the experiment should be conducted in four steps. The first step is to do a sample trial of the experiment to verify the method. The second step is to conduct the full experiment in a garrison environment, followed by an analysis of the compiled data, and presentation of the result. The third step is to determine if the field environment trial of the experiment should be conducted. If so, the fourth step is to conduct the field environment trial, analyze the compiled data, and present the field environment data along with the comparison analysis. This last step is optional as was discussed in Chapter Four.

Also discussed in this chapter are the preparation, a sample computer run, and the steps in the execution of the experiment. The final subject of the chapter is a summary of the costs of conducting the full experiment.

## PREPERATION FOR THE EXPERIMENT

The presentation of the material in Chapter Three was done with the assistance and criticism of individuals from the US Army Institute for the Behavioral and Social Sciences, the US Army Ordnance School, and the US Army Logistics Center. In spite of this, it would be necessary to review the evaluation of the concept and development of the process model, the error classification scheme, and the set of contributing factors. This is done best with the assistance of those individuals selected to be the panel of experts. Once concurrence is acheived with regard to the concept, the process would continue.

The recommendation made in this thesis is that the US Army Research Institute be the organization responsible for directing the experiment. It could be supported by the US Army Ordnance School in providing observers. Additionally, an appropriate agency needs to be identified to assist in providing information to permit selection of and coordination with the Divisions that would provide the sample of mechanics for the garrison trial and the optional field trial. Of course, it is presumed that the experiment could be turned over to a contractor instead of conducting the experiment with US Army assets. If this is done, it is recommended that the US Army Research Institute be the contracting agency.

## ANALYSIS OF EXAMPLE DATA

An example data set was developed to test the statistical procedures and to insure the feasibility of the statistical methods discussed in this thesis. This data contains no preconcieved desire to cause a particular statistical result. It is merely a marginally random distribution of numbers to facilitate a computer test.

The statistical package SPSS$^X$ (Statistics Package for the Social Sciences, Extended) was used, as implemented on the Florida Institute of Technology's DEC VAX 11/780. The code written to support the run is taken from the SPSS$^X$ User's Manual.

## THE SAMPLE TRIAL

A sample trial must be conducted to debug the concept for the experiment, to provide information regarding the training of the observers, and to prove the means of information gathering for the Observer's Recording Form. The sample trial may provide the necessary validation of the above items, or it may be necessary to redo some or all of the preperation. Of course, there is also the possibility that the sample trial may prove the concept to be infeasible. In any case, performing the garrison trial without total confidence in the concept and the method would be a serious error.

A list of subjects to be discussed with the observers is provided at Appendix B. This list contains many of the important topics that the observers should be aware of. As was mentioned in Chapter Four, it is essential that the observer personnel remain as objective as possible. The observers must be trained to be as factual and as non-judgemental as possible. Their task is to observe the mechanic, limit conversation to the minimum necessary to understand the mechanic's purposes and actions, and to record this information along with observed factors which might contribute to possible incorrect procedural activity of the mechanic. Also contained in Appendix B is a sample Observer's Recording Form.

The observers must be active in determining the state of the workplace, the availability of tools, manuals, and diagnostic equipment. They should understand the process model, the error types, and the contributing factors to the degree necessary for their accurate recording of information to be used by the expert panel.

This sample trial will also serve to lend some insight into the average number of errors per mechanic's diagnostic and repair task, and the number of diagnostic and repair tasks per day. This will aid in confirming the number of sample mechanics necessary to obtain the required error sample. There is no doubt that these error rates are subject to great variation depending on the type of unit the mechanic is assigned to and the mission of that unit.

It is not likely that the errors obtained from the sample trial will provide any statistical insight, however, this possibility should not be overlooked. Any insight gained this early may provide an opportunity to strengthen the concept or the method.

## THE GARRISON AND FIELD TRIALS

Following the sample trial, a decision will be made that will determine whether or not the concept and the method are feasible. If both have acheived feasibility, the garrison trial will subsequently be conducted. The reason for conducting the garrison trial first is to take advantage of further gains in experience acheived in this trial prior to conducting what is likely to be the more demanding field trial. Additionally, regardless of the demands of command personnel, Army companies, troops, and batteries tend to be highly selective in the material and equipment that is taken from the garrison maintenance shop to the field due to the transitory nature of field exercises. Units ordinarily take what is essential for survival for the duration of the field exercise, with many items left behind to be used only in garrison. Although it is of interest to know the impact of such expediency, this should be done only after the garrison trial has determined the scope of the problem under the unit's ideal operating environment. Then the real significance of the field trial would lie in the discovery of reduced effectiveness of mechanics in a field environment and whether this is the mechanic's fault or if it is due to expediency on the part of the company, troop, or battery.

The garrison trial will provide the best framework for a thorough evaluation of all elements of the US Army system that touch on the success of the mechanic in accomplishing his mission. The field environment may reduce the quality of this evaluation such that a misleading result would be obtained with no means for comparison, had the field trial been done first and no garrison trial done subsequently.

It is the recommendation of the author, that a positive decision following the sample trial should result in the conduct of the garrison trial first. Upon successful analysis and presentation of the garrison trial data, and a decision to proceed, the field trial would then be conducted. The field trial analysis would be completed and followed by comparing the data of the two trials or treatments to discover the nature and significance of any differences.

## SELECTING THE SAMPLE

The sample mechanics for this experiment will be selected from the entry level of Military Occupational Specialties which are designated as Organizational Level Vehicle or Automotive Mechanics. These specialties are listed in Table 5-1. These specialties are not all used in every type of Division in the US Army. They do represent those mechanics who perform

Table 5-1

Military Occupational Specialties to be Used in the Experiment

| MOS | Description |
|---|---|
| 63B10 | Light Wheeled Vehicle Mechanic |
| 63D10 | Self-Propelled Artillery Automotive Mechanic |
| 63E10 | M1 Tank Automotive Mechanic |
| 63N10 | M60 Tank Automotive Mechanic |
| 63S10 | Heavy Wheeled Vehicle Mechanic |
| 63T10 | ITV/IFV/CFV Automotive Mechanic |
| 63Y10 | Other Tracked Vehicle Automotive Mechanic |

the bulk of diagnosis and repair of wheeled and tracked automotive type vehicles at organizational level. It is these mechanics who are the subject of this experiment.

A random selection process must be performed at two levels. The first selection is relatively easy, it involves the selection of five Divisions from among the active duty Divisions that may be identified as accessible for experimentation by the Chief of Staff of the Army. The second level of selection is more time consuming. It involves direct contact with the Adjutant of each Division to obtain the list of soldiers from within that Division who possess the required specialties at the entry skill level. These names can be obtained from the Division's automated personnel roster. The soldier's identification numbers are then used in a random selection process to identify a list of five primary and ten alternate selections. Coordination is then made through the Division Adjutant to verify the availability of the primary selectees during a predetermined observation period. Substitutions are made as necessary until five mechanics have been identified who will be working in garrison during the planned observation period.

The field trial selection process will be accomplished in identical fashion. It will include Division selection, mechanic selection, and verification of mechanic availability in a field setting. For the field trial, observation periods would need to be coordinated to coincide with major field training exercises, or perhaps to be done at the National Training Center.

The sample trial need not require rigorous sampling techniques as it is only a test of the concept and method. Observation of two mechanics with one of the appropriate specialties for a period of one week, is all that would be necessary. It is possible, of course, that the sample trial may need to be repeated to insure adequate confirmation of the concept and the method.

## SUMMARY OF COSTS

There are costs associated with the sample trial, the garrison trial, and the field trial. The costs to accomplish the sample trial are relatively modest. The costs for the garrison and field trials would be nearly identical save for the extended analysis following the field trial. This extended analysis would be necessary to determine the nature and significance of the "between-treatment" differences. Table 5-2 shows the costs associated with each of the trials. Appendix C contains the explanation of how these figures were derived. Also included in Appendix C is a discussion of the personnel requirements for the conduct of each trial.

Table 5-4

Total Costs

| Item | Cost |
|---|---|
| Sample Trial | |
| 1. Personnel | $9582.00 |
| 2. Travel | 750.00 |
| 3. Per Diem | 3150.00 |
| Garrison Trial | |
| 1. Personnel | 39,177.50 |
| 2. Travel | 15,600.00 |
| 3. Per Diem | 22,050.00 |
| 4. Other | 1100.00 |
| Field Trial | |
| 1. Personnel | 39177.50 |
| 2. Travel | 15,600.00 |
| 3. Per Diem | 22,050.00 |
| 4. Other | 1100.00 |
| Total | 167,337.00 |

## Chapter 6

## CONCLUSION

In closing, it is important to re-examine the purpose of this thesis. Given that there is currently no means to optimize the allocation of research and development funds to improve the performance of soldier mechanics in the US Army; there is a need to develop a method to provide the necessary insight for optimization. This thesis has developed a concept and a method that should be considered in the effort to provide useful information to US Army decision makers.

Although the methods of Human Factors that involve analysis of human errors have been demonstrated and documented, this is a relatively new area. This thesis proposes an application of error analysis that is similar to what has been documented, Van Eekhout and Rouse [81], Rouse and Rouse [83]. However, it is felt that the concept of correlating error types and contributing factors to reach causes must withstand further scrutiny by individuals in the fields of Human Factors and Research Psychology prior to conducting the proposed experiment. The list of error types and the development of contributing factors requires further refinement.

It is hoped that this thesis will proivide a basis for further discussion of methods to optimize the allocation of funds to improve the performance of soldier mechanics.

# BIBLIOGRAPHY


Buchan, Ross S., Knutson, Ray. "The ITDT Program, the Army Departs from Tradition," Defense Management Journal. January 1977, pp. 33-37.

Dressel, Douglas J., Shields, Joyce L. Organizational Maintenance Performance. US Army Research Institute for thr Behavioral and Social Sciences, Research Problem Review 79-8, April 1979.

Federico, Pat-Anthony. Accomodating Instruction to Student Characteristics: Trends and Issues. US Navy Personnel Research and Development Center, Technical Report 79-1, October 1978.

Hackenbruch, Douglas J. "Fixing Forward Needs Fixing," Army Logistician. November/December 1982, pp. 29-31.

Harris, Douglas H., Fuller, R. G., Harper, W. R. The Maintenance Performance System (Organizational) Description of Components and Operation. Santa Barbara, CA., ANACAPA Sciences Incorporated, for US Army Research Institute for the Behavioral and Social Sciences, Technical Report 465-34, May 1984.

Janis, I. L., Mann, L. A Psychological Analysis of Conflict, Choice, and Commitment. New York: The Free Press, 1977.

Johnson, William B., Rouse, William B. "Analysis and Classification of Human Errors in Troubleshooting Live Aircraft Power Plants," IEEE Transactions on Systems, Man, and Cybernetics, SMC-XII, No. 3, May/June 1982, pp. 389-393.

Kern, Richard P., Hayes John F. Research Findings to Aid Supervisors and Trainers in Improving Maintenance Performance. US Army Research Institute for the Behavioral and Social Sciences, Research Product 83-14, November 1983.

Kern, Richard P. "Modeling Users and Their Use of Technical Manuals (Draft)," Unpublished, Undated

Lindeman,Richard H., Merenda, Peter F. Introduction to Bivariate and Multivariate Analysis. Scott, Foresman and Company, Glenview, IL.,1980.

Pantiskas, Carl A. "TMDE, Present and Future," Army Logistician, January/February 1981, pp. 30-32.

Rasmussen, Jens, Rouse, William B. eds., Human Detection and Diagnosis of System Failures. Plenum, New York, 1980.

Rouse, Sandra H., Rouse, William B., "Cognitive Style as a Correlate of Human Problem Solving Performance in Fault Diagnosis Tasks," IEEE Transactions on Systems, Man, and Cybernetics, SMC-XII, No. 5, September/ October 1982, pp. 649-652.

Rouse, William B., Rouse Sandra H., "Analysis and Classification of Human Error," IEEE Transactions on Systems, Man, and Cybernetics, SMC-XIII, No. 4, July/August 1983, pp. 539-549.

Rouse, William B., "A Model of Human Decision Making in Fault Diagnosis Tasks that Include Feedback and Redundancy," IEEE Transactions on Systems, Man, and Cybernetics, SMC-IX, No. 4, April 1979A, pp. 237-241.

Rouse, William B., Rouse, Sandra H. "Measures of Complexity of Fault Diagnosis Tasks," IEEE Transactions on Systems, Man, and Cybernetics, SMC-IX, No. 11, November 1979B, pp. 720-727.

Sage, Andrew P. "Behavioral and Organizational Cosiderations in the Design of Information Systems and Processes for Planning and Decision Support," IEEE Transactions on Systems, Man, and Cybernetics, SMC-XI, No. 9, September 1981, pp. 640-678.

Schurman, D. L., et al, Guidelines: Assessing Use of Information Sources and Quality of Performance at the Work Site. US Army Institute for the Behavioral and Social Sciences, Research Note 82-7, December 1980.

Staff, "CMF 63 Changed," Army Logistician. July/August 1980, pp. 22-25.

Staff, "CMF 63 Changes Proposed," Army Logistician. September/November 1978, pp. 22-25.

Stefik, M. et al, "The Organization of Expert Systems, A Tutorial," Artificial Intelligence Journal. XVIII, No. 2, March 1982, pp. 130-140.

Sweeny, Bruce D. "Forward Area TMDE," Army Logistician. September/October 1981, pp. 14-15.

US Department of the Army, DA Board of Inquiry on the Army Logistics System. 1966.

US Department of the Army, Operational Concept for Unit and Intermediate Maintenance Diagnostician. US Army Ordnance Center and School, May 1984.

US Department of the Army, Program of Instruction for MOS 63B10, Light Wheel Vehicle Mechanic (Draft). US Army Ordnance Center and School, Aberdeen Proving Ground, MD., April 1984.

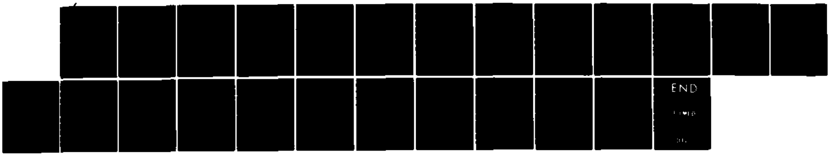
US Department of the Army. Troubleshooting Test Conducted by the USAMMCS. US Army, Fort Carson, Colorado, 1-31 July 1974.

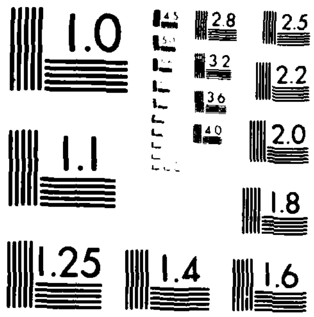
Van Eekhout, Joost M., Rouse, William B., "Human Errors in Detection, Diagnosis, and Compensation for Failures in the Engine Control Room of a Supertanker," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-11, No. 12, December 1981, pp. 813-816.

Wohl, Joseph G. "Cognitive Capability Versus System Complexity in Electronic Maintenance," IEEE Transactions on Systems, Man, and Cybernetics, SMC-XIII, No. 4, July/August 1983, pp. 624-626.

AD-A152 229 THE DESIGN OF AN EXPERIMENT TO EXAMINE REPAIR PROCESS 2/2
ERRORS OF MILITARY VEHICLE MECHANICS(U) FLORIDA INST OF
TECH MELBOURNE D T CLEMENTS 07 DEC 84

UNCLASSIFIED F/G 5/9 NL






MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS

## APPENDIX A

## THE SAMPLE COMPUTER RUN

The sample computer run was accomplished with the reduced example data-set depicted in Table A-1. The independent variable set X1 through X7 are the contributing factors. The dependent variables, Y1 through Y7, are the error types.

### The Test for Homogeneity

The test for homogeneity used here compares the variability of the observed major error category totals of each mechanic with the mean of all of the mechanics. Also included is the same comparison for total errors. The Chi-Squared statistic is calculated as follows:

$$\mathrm{Chi}_j^2 = \sum_{i=1}^{8} (O_{ij} - E_j)^2 / E_j$$

where:

$O_{ij}$ = observed cell frequency for

error category i $i = 1, \ldots, 8$

mechanic j $j = 1, \ldots, 20$

$E_i$ = expected value of error category i

$Chi_j^2$ = Chi-Squared statistic for mechanic j

$df = n-1 = 8-1 = 7$

Table A-2 shows the cell measurements and the Chi-Squared statistics for each mechanic. If the test is used to eliminate the 5% tail, two mechanics are rejected, if the 10% tail is removed, six mechanics are rejected.

Table A-1

Example Data

| Mech | X1 | X2 | X3 | X4 | X5 | X6 | X7 | Y1 | Y2 | Y3 | Y4 | Y5 | Y6 | Y7 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 12 | 11 | 8 | 6 | 15 | 13 | 9 | 6 | 8 | 3 | 18 | 2 | 30 | 7 |
| 2 | 8 | 11 | 4 | 10 | 7 | 5 | 7 | 4 | 10 | 2 | 17 | 5 | 14 | 0 |
| 3 | 7 | 12 | 14 | 6 | 9 | 10 | 9 | 5 | 11 | 1 | 15 | 4 | 20 | 11 |
| 4 | 11 | 8 | 6 | 14 | 9 | 10 | 15 | 3 | 15 | 0 | 17 | 3 | 30 | 5 |
| 5 | 13 | 10 | 14 | 16 | 18 | 9 | 13 | 2 | 14 | 5 | 16 | 2 | 34 | 10 |
| 6 | 8 | 9 | 6 | 12 | 11 | 5 | 8 | 4 | 12 | 4 | 15 | 5 | 15 | 4 |
| 7 | 12 | 9 | 17 | 9 | 11 | 12 | 18 | 4 | 11 | 2 | 14 | 2 | 40 | 15 |
| 8 | 7 | 14 | 8 | 12 | 12 | 8 | 12 | 7 | 10 | 1 | 13 | 17 | 25 | 10 |
| 9 | 9 | 6 | 13 | 12 | 8 | 6 | 9 | 3 | 8 | 7 | 17 | 3 | 18 | 7 |
| 10 | 11 | 14 | 16 | 12 | 10 | 15 | 14 | 2 | 12 | 8 | 14 | 2 | 40 | 15 |
| 11 | 6 | 11 | 5 | 9 | 8 | 6 | 9 | 5 | 14 | 3 | 13 | 2 | 15 | 2 |
| 12 | 8 | 6 | 9 | 7 | 14 | 8 | 10 | 4 | 11 | 2 | 15 | 5 | 18 | 7 |
| 13 | 12 | 14 | 6 | 9 | 11 | 7 | 12 | 3 | 13 | 1 | 16 | 4 | 24 | 10 |
| 14 | 13 | 6 | 15 | 11 | 14 | 14 | 15 | 2 | 14 | 7 | 18 | 3 | 34 | 10 |
| 15 | 14 | 7 | 8 | 11 | 14 | 7 | 8 | 6 | 10 | 6 | 19 | 2 | 16 | 10 |
| 16 | 9 | 14 | 7 | 11 | 15 | 8 | 14 | 8 | 13 | 5 | 18 | 5 | 19 | 10 |
| 17 | 14 | 7 | 12 | 19 | 8 | 12 | 13 | 1 | 12 | 2 | 14 | 6 | 40 | 10 |
| 18 | 10 | 8 | 11 | 6 | 13 | 6 | 8 | 3 | 10 | 1 | 15 | 3 | 25 | 5 |
| 19 | 13 | 17 | 10 | 11 | 11 | 11 | 17 | 2 | 15 | 3 | 17 | 2 | 41 | 10 |
| 20 | 9 | 1 | 11 | 16 | 8 | 10 | 16 | 1 | 17 | 4 | 14 | 4 | 31 | 10 |

Table A-2

Test For Homogeneity

| case | 1 | 2 | 3 | 4 | 5 | 6 | 7 | tot | $Chi_j^2$ | .10 | .05 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 12 | 11 | 8 | 6 | 15 | 13 | 9 | 74 | 9.70 | | |
| 2 | 8 | 11 | 4 | 10 | 7 | 5 | 7 | 52 | 16.09 | X | X |
| 3 | 7 | 12 | 14 | 6 | 9 | 10 | 9 | 67 | 7.09 | | |
| 4 | 11 | 8 | 6 | 14 | 9 | 10 | 15 | 73 | 4.43 | | |
| 5 | 13 | 10 | 14 | 16 | 18 | 9 | 13 | 93 | 13.76 | X | |
| 6 | 8 | 9 | 6 | 12 | 11 | 5 | 8 | 59 | 8.39 | | |
| 7 | 12 | 9 | 17 | 9 | 11 | 12 | 18 | 88 | 12.62 | X | |
| 8 | 7 | 14 | 8 | 12 | 12 | 8 | 12 | 73 | 3.11 | | |
| 9 | 9 | 6 | 13 | 12 | 8 | 6 | 9 | 63 | 7.17 | | |
| 10 | 11 | 14 | 16 | 12 | 10 | 15 | 14 | 93 | 14.53 | X | X |
| 11 | 6 | 11 | 5 | 9 | 8 | 6 | 9 | 54 | 13.37 | X | |
| 12 | 8 | 6 | 9 | 7 | 14 | 8 | 10 | 62 | 6.71 | | |
| 13 | 12 | 14 | 6 | 9 | 11 | 7 | 12 | 71 | 4.19 | | |
| 14 | 13 | 6 | 15 | 11 | 14 | 14 | 15 | 88 | 11.81 | | |
| 15 | 14 | 7 | 8 | 11 | 14 | 7 | 8 | 69 | 5.42 | | |
| 16 | 9 | 14 | 7 | 11 | 15 | 8 | 14 | 78 | 6.62 | | |
| 17 | 14 | 7 | 12 | 19 | 8 | 12 | 13 | 85 | 12.62 | X | |
| 18 | 10 | 8 | 11 | 6 | 13 | 6 | 8 | 62 | 7.25 | | |
| 19 | 13 | 17 | 10 | 11 | 11 | 11 | 17 | 90 | 11.44 | | |
| 20 | 9 | 11 | 11 | 16 | 8 | 10 | 16 | 81 | 6.59 | | |
| tot | 206 | 205 | 200 | 219 | 226 | 182 | 236 | 1474 | | | |

E(x) 1,10.30 2,10.25 3,10.00 4,10.95 5,11.30 6,9.10 7,11.80 tot,73.75

$Chi^2_{7,0.10} = 12.017$

$Chi^2_{7,0.05} = 14.067$

The X's indicate the mechanics to be eliminated for each alpha.

## The Factor Analysis

The factor analysis of the variables X1 through X7 provided no reduction in the variable set. The factor analysis of the variables Y1 through Y7 did provide a reduction to three factors as shown in Table A-3. It would be reasonable to use this reduced set of error-type factors in a canonical correlation inasmuch as discrimination is required among the contributing factors, not the error types. If a reduction of the set of X's is acheived, an examination of the new factors must be made to determine if the new factor subsets cut across the boundaries of the represented primary contributing factors as described in Chapter 3. If so, the factor analysis of the X's may have to be abandoned as there would be no way to distinguish the systemic weaknesses that may be identified in the canonical correlation.

Table A-3

Factor Analysis*

| Variable | Factor 1 | Factor 2 | Factor 3 |
|---|---|---|---|
| Y6 | 0.83266 | | |
| Y7 | 0.80463 | | 0.35639 |
| Y1 | 0.63097 | | |
| Y3 | 0.61695 | | |
| Y4 | 0.50109 | | |
| Y5 | | -0.84148 | |
| Y2 | | | 0.65542 |

*Using SPSS$^{X}$ Principal Axis Factoring

Canonical Correlation

The calculations for canonical correlation analysis for the example data-set will now be described per Lindeman [80]. Canonical correlation will relate the set of X's to the set of Y's by assigning linear weights (similar to regression coefficients) to each variable in the set. These linear weights describe the relative importance of each variable to the relationship. The variable weight value is not significant of itself, only its relative size is significant. As an example, a weight value of 0.9 is three times as significant as a value of 0.3, the value 0.9. of itself, says nothing. This relationship is described as:

$$\alpha_1 Y_1 + \ldots + \alpha_p Y_p = \beta_1 X_1 + \ldots + \beta_q X_q$$

Let:

$$Z_y = \alpha_1 Y_1 + \ldots + \alpha_p Y_p$$

$$Z_x = \beta_1 X_1 + \ldots + \beta_q X_q$$

The Z's are called the canonical variates. As in any analysis of variance, there is a correlation coefficient, here called the canonical correlation coefficient, that explicitly describes the quality of the relationship between the X's and the Y's, by stating the percent of variation explained by the variates, one for the other. The canonical correlation coefficient is $R_c$, and the percent explained variation value is obtained as follows:

$$\%\text{ Explained Variation} = 100(R_c^{\,2})$$

The calculations for canonical correlation are as follows:

Let:

$$M = R_{yy}^{-1} R_{yx} R_{xx}^{-1} R_{xy}$$

where the R's are from the ordinary unit variance correlation matrix that is set up X1 to X7,Y1 to Y7 across the top, by X1 to X7, Y1 to Y7 down the side.

Solve:

$$|M-\lambda_j| = 0$$

This calculation will produce a polynomial equation in $\lambda$. Solve for the positive roots of the $\lambda$'s, these are the eigenvalues of the matrix M. There may be as many non-zero $\lambda$'s as there are variables in the smaller variable set Y or X.

To determine the significance of the $\lambda$'s and consequently to test the experimental null hypothesis that there is no correlation between the set of X's and the set of Y's, the following is used:

$$H_0: \lambda_1 = \lambda_2 = . . . = \lambda_j = 0$$

Calculate the root k for k = 1, . . . , j

$$\Lambda_k = \prod_{j=k}^{r} (1-\lambda_j) \qquad \text{(Due to Wilk)}$$

where:

$\Lambda_k$ Wilk's Lambda for root k

$r = \min(p,q)$

Calculate:

$$V_k = -[n-1-1/2(p+q+1)]\ln\Lambda_k \qquad \text{(Due to Bartlett)}$$

where:

$V_k$ is Chi-Squared distributed

with:

$d.f. = (p-k+1)(q-k+1)$ for large n

$n = \max(100,20(p+q))$

Reject $H_0$ when significance of statistic for $\lambda_j$ is beyond 0.05, Lindeman [80].

Continuing with the calculations for canonical correlation, it should be noted that the maximum number of non-zero roots is equal to min(p,q). One or more of these roots may be significant. In this correlation, however, only root $\lambda_1$ is of interest. The canonical correlation coefficient:

$$R_{c1} = \lambda_1^{1/2}$$

The next step is to obtain the variates $Z_y$ and $Z_x$. These variates are vectors. $Z_{y1}$ is the eigenvector $\mathbf{d}_1$ and $Z_{x1}$ is the eigenvector $\mathbf{c}_1$, both obtained from the first eigenvalue $\lambda_1$. *As can be seen, there is a pair of* variates $Z_{yj}$, $Z_{xj}$ and a canonical correlation coefficient $R_{cj}$ for each non-zero root. As only $\lambda_1$ is of interest, the subscripts for root identification will now be assumed to be 1 and dropped from the notation.

Any column from adj(M-$\lambda$I) is an eigenvector associated with $\lambda$. Select any one column and standardize it as follows:

Let $f^t$ (f transpose) be the cofactors of the first row of M-$\lambda$I and let:

$$\theta = (f^t R_{yy} f)^{1/2}$$

and

$$\mathbf{d} = (1/\Theta)\mathbf{f}$$

The vector **d** contains the weights associated with the set of Y's. Now calculate the companion vector **c** as follows:

$$\mathbf{c} = \lambda^{-1/2} R_{xx}^{-1} R_{xy}\,\mathbf{d}$$

The vector **c** contains the weights associated with the set of X's. The **d** is the canonical variate $Z_y$ and the vector **c** is the canonical variate $Z_x$.

The statistical package SPSS$^X$ was used on the example data-set. The following results were obtained:

$$\lambda_1 = 0.99562$$

$$R_{c1} = 0.99781$$

This means that 99.562% of the variation in $Z_{y1}$ is explained by $Z_{x1}$ and vice versa.

Performing the test of significance on $\lambda_1$ yeilds the following result:

$$\Lambda_1 = 0.00008599$$

$$V_1 = 13718.887$$

$$\text{d.f.} = 49$$

$$\text{Chi}^2_{49} = 13718.887$$

This statistic is significant well beyond 0.05, and consequently $H_0: \lambda_1 = 0$ is rejected.

$$
\begin{aligned}
Z_{y1} = \mathbf{d}_1 = \; & 0.13156 && = \alpha_1 \\
& 0.06407 && = \alpha_2 \\
& 0.05981 && = \alpha_3 \\
& 0.02105 && = \alpha_4 \\
& 0.02172 && = \alpha_5 \\
& 0.09028 && = \alpha_6 \\
& 0.10060 && = \alpha_7
\end{aligned}
$$

$$
\begin{aligned}
Z_{x1} = \mathbf{c}_1 = \; & 0.04357 && = \beta_1 \\
& 0.06917 && = \beta_2 \\
& 0.02885 && = \beta_3 \\
& 0.10702 && = \beta_4 \\
& -0.00009 && = \beta_5 \\
& 0.07811 && = \beta_6 \\
& 0.08383 && = \beta_7
\end{aligned}
$$

These vector values or weights are then used to rank order the variables. Table A-4 and Table A-5 show how the information can be used to produce decision aids for the decision-maker.

Table A-2

Factors Which Contribute to Mechanic's Errors

| Rank | Variable | % of Errors Associated with This Variable | Coefficient |
|---|---|---|---|
| 1 | X4 | 21% | 0.10702 |
| 2 | X7 | 11% | 0.08383 |
| 3 | X6 | 36% | 0.07811 |
| 4 | X2 | 16% | 0.06917 |
| 5 | X1 | 5% | 0.04357 |
| 6 | X3 | 5% | 0.02885 |
| 7 | X5 | 5% | -0.00009 |

Table A-3

Types of Mechanic's Errors Most Likely to be Reduced
by the Mitigation of Contributing Factor Influence

| Rank | Variable | % of Total Errors | Coefficient |
|---|---|---|---|
| 1 | Y1 | 14% | 0.13156 |
| 2 | Y7 | 16% | 0.10060 |
| 3 | Y6 | 12% | 0.09028 |
| 4 | Y2 | 14% | 0.06407 |
| 5 | Y3 | 14% | 0.05981 |
| 6 | Y5 | 15% | 0.02172 |
| 7 | Y4 | 15% | 0.02105 |

## APPENDIX B

## INSTRUCTIONS TO THE OBSERVERS

An excellent guide to the training and use of observers in an experiment similar to this thesis is contained in Schurman et al [80].

Initial Briefing

1. The purpose of the observation is to determine how closely mechanics follow the prescribed methods in doing their tasks.
2. Explain the Process Model.
3. Explain the Error Classification Scheme.
4. Explain the Contributing Factors.
5. Explain the Observer's Recording Form and how to use it.

Prior to Arrival at the Mechanic's Unit

1. Know the Military Occupational Specialty of the mechanic.
2. Read the Soldier's Manual for the Specialty.
3. Review the Program of Instruction for the Specialty.

4. Know how long it has been since the mechanic completed Initial Entry Training.

5. Know how long it has been since the mechanic was assigned to this unit.

Upon Arrival at the Mechanic's Unit

1. Overcome any fear or apprehension that may surround the observer's arrival at the unit (company, troop, or battery).

2. The observer team leader and the Division representative will take the observer to the unit and meet with the Unit Commander.

3. Use the previously defined initial briefing to explain the purposes of the observer's work to the Unit Commander.

4. Explain that this is not an evaluation of the unit or the mechanic.

5. *The data from this mechanic will not be used in any way that will* reflect poorly on the mechanic, his unit, or the Division.

6. Brief the Maintenance leader and the mechanic's supervisor.

7. The observer should brief the mechanic in the presence of the mechanic's supervisor.

The Observation Period

1. Actions and communication by the observer should be totally

non-judgemental.

2. Elicit conversation and contact only to the extent necessary to understand the mechanic's mental and physical processes, as they relate to his doing his tasks.
3. Observe the workplace, the mechanic, his supervisor, and the environmental conditions.
4. Do not check on the availability of tools, manuals, or equipment during this period.

<u>The Post Observation Period</u>

1. In those cases where there was a question regarding the availability of tools, manuals, or other equipment, check to see if the items are present.
2. In those cases where there was a question regarding equipment calibration or currency of manuals, check these items now.
3. Contributing factor identification may change as a result of the above.

<u>The Observer's Recording Form Figure B-1</u>

Block No.

1. Observer identification

2. Mechanic identification

3. Date and time that the task was started

4. Date and time that the task was completed

5. The description of the task, use the Soldier's Manual task identification if possible.

6. The vehicle nomenclature

7. Circle the steps of the Process Model that were used by the mechanic

8. List each discrete step or item of note that was observed during the task performance procedure, number each item

9. List the item numbers of suspected errors, the contributing factors, and explain the contributing factor selection

10. List any other information that may be of interest in the examination of this report

Figure B-1

Observer's Recording Form

---

1. Observer No.__________ 2. Mechanic No.__________
3. Date/Time Started__________ 4. Date/Time Completed__________
5. Task Description__________
6. Vehicle Nomenclature__________

7. Process Steps 1 2 3 4 5 6 7

8. List Each Item of Task Performance Observed, Number Each Item

9. Item numbers of suspected errors with contributing factors, use reverse
10. List any other information of interest to the expert panel, use reverse

---

## APPENDIX C

## PERSONNEL REQUIREMENTS AND COSTS

This appendix will develop the personnel requirement for each of the experimental trials. Also to be explained are the costs for personnel travel, per diem for temporary duty, and miscellaneous costs.

### The Sample Trial

The sample trial team should include a Research Psychologist as the leader, who would also act as a one-person expert-panel. At least two observers, experts in vehicle maintenance, should be used to take advantage of any differences of opinion that exist between the two regarding the trial circumstances. A statistical analyst from the field of Research Psychology should be included to confirm the statistical methods. Lastly, an advisor from the field of vehicle maintenance should be included to comment on the overall concept and its utility. This advisor should be the person who will make the decision to go on with the garrison trial, or he should be the decision-maker's representative

## The Garrison Trial

The team leadership for the garrison trial should consist of the decision-maker's representative, and a Research Psychologist. The panel of two or more experts in vehicle maintenance to score the error recording forms would work under the team leaders. The statistical analyst would also work under the team leader. Ideally, the observer team would consist of five observers with a sixth individual as leader. The leader would provide liaison with the Divisions supplying the sample mechanics. The observers should be kept free from the logistical concerns such as travel and lodging so that their time is spent with the mechanics.

## The Field Trial

The field trial team should be organized in the same manner and size as the garrison trial team. The garrison team leader, the expert panel, and the analyst should be retained for continuity and to take advantage of their garrison trial experience. Although a similar argument may exist for the observer team, the concern regarding observer bias must be considered, and will likely favor changing the observers for the field trial. This new observer team will also re-inforce needed independence between replications.

## Conclusion

This concludes the summary of personnel requirements. Tables C-1 and C-2 list the likely costs for the personnel, given one calendar week per observed sample mechanic , and twenty-five sample mechanics. The additional five mechanics provide the buffer necessary to protect the sample size from losses due to homogeneity test rejections.

The travel costs are for an average of round-trip airfares from Washington DC to continental US posts, the Republic of Korea, and the Federal Republic of Germany. The travel to these locations reflects the needs of the random process of selection of Divisions. Temporary Duty pay is calculated at the maximum rate in all cases. The salary rates reflect the pay of middle and upper grade civilian personnel. Office supplies include the printing of recording forms and miscellaneous items. Computer costs are estimates of CPU time based on running the example data. Costs are based upon published schedules in effect during November 1984.

Table C-1

The Sample Trial Costs

| Item | Cost |
|---|---|
| Personnel (Salaries) | |
| 1. Leader/Expert GS 15 | |
| 4 weeks at $1150/week | $4600.00 |
| 2. Analyst GS 14 | |
| 4 weeks at $958/week | 3832.00 |
| 3. Observers (2) GS11 | |
| 2 weeks at 287.50/week | 1150.00 |
| Travel (Stateside Only) | |
| 1. Leader/Expert | 250.00 |
| 2. Observers (2) | 500.00 |
| (Example used Ft. Carson, CO) | |
| Per Diem | |
| 1. Leader/Expert | 1050.00 |
| 2. Observers (2) | 2100.00 |
| Total | $13,482.00 |

Table C-2

Garrison and Field Trial Costs

| Item | Cost |
|---|---|
| Personnel (Salaries) | |
| 1. Research Psychologist GS15 | |
| 4 weeks at $1150/week | $4600.00 |
| 2. Maintenance Expert GS 14 | |
| 4 weeks at $958/week | 3832.00 |
| 3. Analyst GS 14 | |
| 4 weeks at $958/week | 3832.00 |
| 4. Experts for Panel (3) GS 14 | |
| 4 weeks at $958/week | 11,496.00 |
| 5. Observer Team Leader GS 13 | |
| 7 weeks at $ 765/week | 5355.00 |
| 6. Observers (5) GS 11 | |
| 7 weeks at $287.50/week | 10,062.50 |
| Travel (Average per text $520.00 each) | |
| 1. Observer Team Leader 5 trips | 2600.00 |
| 2. Observers (5) 5 trips | 13,000.00 |
| Per Diem (use 75.00 per day) | |
| 1. Observer Team Leader | |
| 49 days | 3675.00 |
| 2. Observers (5) | |
| 49 days | 18,375.00 |
| Other | |
| 1. CPU Time ( $100.00 per hour) | 100.00 |
| 1 hour | |
| 2. Office Supplies | 1000.00 |
| Total | 76,927.50 |

END

FILMED

5-85

DTIC